

Swope, Sheridan

From: Swope, Sheridan
Sent: Thursday, August 25, 2005 11:48 AM
To: Sheppard, Paula
Subject: FW: 10/726,967

what's going on with this search?

-----Original Message-----

From: Swope, Sheridan
Sent: Wednesday, August 10, 2005 8:42 PM
To: STIC-Biotech/ChemLib
Subject: 10/726,967

For 10/726,967, pls search:

SID 57 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Barb O'Brien

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2005, 00:28:26 ; Search time 333 Seconds
(without alignments)
8.130 Million cell updates/sec

Title: US-10-726-967a-57

Perfect score: 35
Sequence: 1 EINHETD 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A_Geneseq_16Dec04:*
2: geneseqp1980s:*
3: geneseqp1990s:*
4: geneseqp2000s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	7	8	ADP83931 Human BAC
2	35	100.0	433	8	ADP83957 Human BAC
3	33	94.3	7	8	ADP83930 Human BAC
4	33	94.3	433	8	ADP83954 Human BAC
5	31	88.6	7	8	ADP83932 Human BAC
6	31	88.6	54	4	ABG23686 Novel hum
7	31	88.6	282	2	AAV35321 Protein i
8	31	88.6	282	6	ABU27126 Protein e
9	31	88.6	282	6	ABU26773 Protein i
10	31	88.6	287	2	AAV37496 Protein i
11	31	88.6	495	4	AAW79357 Human pro
12	31	88.6	887	4	ABG23684 Novel hum
13	31	88.6	1113	4	ABG06702 Novel hum
14	30	85.7	146	5	ADK34827 Novel hum
15	30	85.7	272	8	ADP08273 Neisseria
16	30	85.7	333	4	AA667225 Amino aci
17	30	85.7	403	8	ADN18450 Bacterial
18	30	85.7	437	6	ABP78779 N. gonorr
19	30	85.7	437	8	ADP08287 Neisseria
20	30	85.7	506	7	ADP37241 Plant yie
21	30	85.7	506	8	AD141903 Plant pro
22	30	85.7	767	6	ADA54801 Human pro
23	30	85.7	929	8	ADS22231 Bacterial
24	30	85.7	958	6	ABU41216 Protein e
25	30	85.7	959	6	ABM67228 Phototrab

26	30	85.7	964	7	ADP04750 Bacterial
27	30	85.7	1400	6	ABP55412 Human MDD
28	30	85.7	1987	4	ABB61083 Drosophi
29	30	85.7	4345	4	ABB66417 Drosophi
30	29	82.9	75	5	ABBS7438 Human sec
31	29	82.9	85	4	AAU36722 Staphyloc
32	29	82.9	96	5	ABP30778 Streptoco
33	29	82.9	110	6	ABP80107 N. gonorr
34	29	82.9	110	6	ABP78140 N. gonorr
35	29	82.9	128	4	ABG02051 Novel hum
36	29	82.9	167	2	ABG02049 Novel hum
37	29	82.9	176	2	AAV10968 H. pylori
38	29	82.9	184	2	AAW20635 H. pylori
39	29	82.9	208	5	ABP40126 Staphyloc
40	29	82.9	208	8	ADSO5107 Staphyloc
41	29	82.9	272	8	ADQ39944 Human myo
42	29	82.9	273	4	AA681829 S. epider
43	29	82.9	273	5	AAU80271 Heterosig
44	29	82.9	275	6	ABU26049 Protein e
45	29	82.9	284	6	ABU37757 Protein e

ALIGNMENTS

RESULT 1
ADP83931
ID ADP83931 standard; peptide; 7 AA.
XX
AC ADP83931;
XX
DT 23-SEP-2004 (first entry)
XX
DE Human BACE1 autoproteolysis site SEQ ID NO:57.
XX
KW human; beta-site amyloid precursor protein cleaving enzyme 1;
KW beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; prodomain;
KW engineered cleavage site; protease domain; neuroprotective; nootropic;
KW gene therapy; Alzheimer's disease; Down's syndrome.
XX
OS Homo sapiens.
XX
PN WO2004056962-A2.
XX
PD 08-JUL-2004.
XX
PF 02-DEC-2003; 2003WO-US038314.
XX
PR 04-DEC-2002; 2002US-0430984P.
XX
PA (SUNE-) SUNESIS PHARM INC.
XX
PI Ballinger M;
XX
DR WPI; 2004-507703/48.
XX
PT New polypeptides for producing homogenously processed preparations of
PT beta site amyloid precursor protein-cleaving enzyme comprises a
PT prodomain, an engineered cleavage site and a protease domain.
XX
PS Claim 10; SEQ ID NO 57; 40pp; English.
XX
The present invention describes a polypeptide (I) comprising in order
from the N-terminus to the C-terminus: (a) a prodomain comprising at
least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID
NO:3 (ADP83877, comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which
is the longest isoform of human beta-site amyloid precursor protein (APP)
cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;
and (c) a protease domain. (I) is capable of being cleaved at the
engineered cleavage site, and so releases a free protease domain that has
BACE1 activity. Also described: (1) a nucleic acid sequence encoding (I);
(2) a vector for expression of (I); and (3) a host cell expressing (I).
(1) has neuroprotective and nootropic activities, and can be used in gene

CC therapy. (1) can be used for producing preparations of homogenously
 CC processed BACE that may be used for e.g. studying or treating diseases
 CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is
 CC located on chromosome 11, more specifically to 11q23.2-23.3. The present
 CC sequence represents a human BACE1 autoproteolysis site amino acid
 CC sequence, which is used in the exemplification of the present invention.
 SQ Sequence 7 AA;

Query Match 100.0%; Score 35; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 7
 |||||
 DB 1 EINLETD 7

RESULT 2
 ADP83957
 ID ADP83957 standard; protein; 433 AA.
 XX
 AC ADP83957;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Human BACE1 mutant amino acid sequence SEQ ID NO:84.
 XX
 KW human; beta-site amyloid precursor protein cleaving enzyme 1;
 KW beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; chromosome 11;
 KW prodomain; engineered cleavage site; protease domain; neuroprotective;
 KW nootropic; gene therapy; Alzheimer's disease; Down's syndrome; mutant.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 WO2004056962-A2.
 XX
 PD 08-JUL-2004.
 XX
 PF 02-DEC-2003; 2003WO-US038314.
 XX
 PR 04-DEC-2002; 2002US-0430984P.
 XX
 PA (SUNE-) SUNESIS PHARM INC.
 XX
 PI Ballinger M;
 XX
 DR WPI; 2004-507703/48.
 XX

PT New polypeptides for producing homogenously processed preparations of
 PT beta site amyloid precursor protein-cleaving enzyme comprises a
 PT prodomain, an engineered cleavage site and a protease domain.
 XX
 PS Example 1; SEQ ID NO 84; 40pp; English.

CC The present invention describes a polypeptide (1) comprising in order
 CC from the N-terminus to the C-terminus: (a) a prodomain comprising at
 CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID
 CC NO:3 (ADP83877, comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which
 CC is the longest isoform of human beta-site amyloid precursor protein (APP)
 CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;
 CC and (c) a protease domain. (1) is capable of being cleaved at the
 CC engineered cleavage site, and so releases a free protease domain that has
 CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (1);
 CC (2) a vector for expression of (1); and (3) a host cell expressing (1).
 CC (1) has neuroprotective and nootropic activities, and can be used in gene
 CC therapy. (1) can be used for producing preparations of homogenously
 CC processed BACE that may be used for e.g. studying or treating diseases
 CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is
 CC located on chromosome 11, more specifically to 11q23.2-23.3. The present
 CC sequence represents a mutant human BACE1 amino acid sequence, which is
 CC used in an example from the present invention.

XX
 SQ Sequence 433 AA;

Query Match 100.0%; Score 35; DB 8; Length 433;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 7
 |||||
 DB 21 EINLETD 27

RESULT 3
 ADP83930
 ID ADP83930 standard; peptide; 7 AA.
 XX
 AC ADP83930;
 XX
 DT 23-SEP-2004 (first entry)
 XX
 DE Human BACE1 autoproteolysis site SEQ ID NO:56.
 XX
 KW human; beta-site amyloid precursor protein cleaving enzyme 1;
 KW beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; prodomain;
 KW engineered cleavage site; protease domain; neuroprotective; nootropic;
 KW gene therapy; Alzheimer's disease; Down's syndrome.
 XX
 OS Homo sapiens.
 XX
 WO2004056962-A2.
 XX
 PD 08-JUL-2004.
 XX
 PF 02-DEC-2003; 2003WO-US038314.
 XX
 PR 04-DEC-2002; 2002US-0430984P.
 XX
 PA (SUNE-) SUNESIS PHARM INC.
 XX
 PI Ballinger M;
 XX
 DR WPI; 2004-507703/48.
 XX

PT New polypeptides for producing homogenously processed preparations of
 PT beta site amyloid precursor protein-cleaving enzyme comprises a
 PT prodomain, an engineered cleavage site and a protease domain.
 XX
 PS Claim 10; SEQ ID NO 56; 40pp; English.

CC The present invention describes a polypeptide (1) comprising in order
 CC from the N-terminus to the C-terminus: (a) a prodomain comprising at
 CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID
 CC NO:3 (ADP83877, comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which
 CC is the longest isoform of human beta-site amyloid precursor protein (APP)
 CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;
 CC and (c) a protease domain. (1) is capable of being cleaved at the
 CC engineered cleavage site, and so releases a free protease domain that has
 CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (1);
 CC (2) a vector for expression of (1); and (3) a host cell expressing (1).
 CC (1) has neuroprotective and nootropic activities, and can be used in gene
 CC therapy. (1) can be used for producing preparations of homogenously
 CC processed BACE that may be used for e.g. studying or treating diseases
 CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is
 CC located on chromosome 11, more specifically to 11q23.2-23.3. The present
 CC sequence represents a human BACE1 autoproteolysis site amino acid
 CC sequence, which is used in the exemplification of the present invention.
 XX

SQ Sequence 7 AA;

Query Match 94.3%; Score 33; DB 8; Length 7;
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINFETD 7
|:|||||
Db 1 EINFETD 7

RESULT 4
ADP83954
ID ADP83954 standard; protein; 433 AA.

AC ADP83954;
XX
DT 23-SEP-2004 (first entry)

DE Human BACE1 mutant amino acid sequence SEQ ID NO:81.

XX human; beta-site amyloid precursor protein cleaving enzyme 1;
KM beta-site App cleaving enzyme 1; BACE1; BACE1 isoform A; chromosome 11;
KM prodomain; engineered cleavage site; protease domain; neuroprotective;
KM neurotropic; gene therapy; Alzheimer's disease; Down's syndrome; mutant.

OS Homo sapiens.
XX Synthetic.

PN W02004056962-A2.

PD 08-JUL-2004.

PF 02-DEC-2003; 2003WO-US038314.

PR 04-DEC-2002; 2002US-0430984P.

PA (SUNE-) SUNESIS PHARM INC.

PI Ballinger M;

DR WPI; 2004-507703/48.

XX New polypeptides for producing homogenously processed preparations of
PT beta site amyloid precursor protein-cleaving enzyme comprises a
PT prodomain, an engineered cleavage site and a protease domain.

PS Example 1; SEQ ID NO 81; 40pp; English.

XX The present invention describes a polypeptide (I) comprising in order
CC from the N-terminus to the C-terminus: (a) a prodomain comprising at
CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID
CC NO:3 (ADP83877, comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which
CC is the longest isoform of human beta-site amyloid precursor protein (APP)
CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;
CC and (c) a protease domain. (I) is capable of being cleaved at the
CC engineered cleavage site, and so releases a free protease domain that has
CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (I);
CC (2) a vector for expression of (I); and (3) a host cell expressing (I).
CC (I) has neuroprotective and neurotropic activities, and can be used in gene
CC therapy. (I) can be used for producing preparations of homogenously
CC processed BACE that may be used for e.g. studying or treating diseases
CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is
CC located on chromosome 11, more specifically to 11q23.2-23.3. The present
CC sequence represents a mutant human BACE1 amino acid sequence, which is
CC used in an example from the present invention.

XX Sequence 433 AA;

Query Match 94.3%; Score 33; DB 8; Length 433;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINFETD 7
|:|||||
Db 21 EINFETD 27

RESULT 5

ADP83932
ID ADP83932 standard; peptide; 7 AA.

AC ADP83932;

DT 23-SEP-2004 (first entry)

DE Human BACE1 autoproteolysis site SEQ ID NO:58.

XX human; beta-site amyloid precursor protein cleaving enzyme 1;
KM beta-site App cleaving enzyme 1; BACE1; BACE1 isoform A; prodomain;
KM engineered cleavage site; protease domain; neuroprotective; neurotropic;
KM gene therapy; Alzheimer's disease; Down's syndrome.

OS Homo sapiens.

PN W02004056962-A2.

PD 08-JUL-2004.

PF 02-DEC-2003; 2003WO-US038314.

PR 04-DEC-2002; 2002US-0430984P.

PA (SUNE-) SUNESIS PHARM INC.

PI Ballinger M;

DR WPI; 2004-507703/48.

XX New polypeptides for producing homogenously processed preparations of
PT beta site amyloid precursor protein-cleaving enzyme comprises a
PT prodomain, an engineered cleavage site and a protease domain.

PS Claim 10; SEQ ID NO 58; 40pp; English.

XX The present invention describes a polypeptide (I) comprising in order
CC from the N-terminus to the C-terminus: (a) a prodomain comprising at
CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID
CC NO:3 (ADP83877, comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which
CC is the longest isoform of human beta-site amyloid precursor protein (APP)
CC cleaving enzyme 1 (BACE1), isoform A); (b) an engineered cleavage site;
CC and (c) a protease domain. (I) is capable of being cleaved at the
CC engineered cleavage site, and so releases a free protease domain that has
CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (I);
CC (2) a vector for expression of (I); and (3) a host cell expressing (I).
CC (I) has neuroprotective and neurotropic activities, and can be used in gene
CC therapy. (I) can be used for producing preparations of homogenously
CC processed BACE that may be used for e.g. studying or treating diseases
CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is
CC located on chromosome 11, more specifically to 11q23.2-23.3. The present
CC sequence represents a human BACE1 autoproteolysis site amino acid
CC sequence, which is used in the exemplification of the present invention.

XX Sequence 7 AA;

Query Match 88.6%; Score 31; DB 8; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EINFETD 7
|:|||||
Db 1 EINFETD 7

RESULT 6
ID ABG23686 standard; protein; 54 AA.

AC ABG23686;

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #23677.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
PA (HYGE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX N-PSDB; AAS87873.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 54045; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 54 AA;

Query Match 88.6%; Score 31; DB 4; Length 54;
Best Local Similarity 71.4%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 7
|:|:|
Db 32 EVNVEETD 38

RESULT 7
AAV35321
ID AAV35321 standard; protein; 282 AA.
XX
XX AAV35321;
XX
XX 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
DE Protein involved in transcription, translation and/or maturation.

XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.
XX
XX Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98MO-IB001890.
XX
PR 21-NOV-1997; 97FR-00014673.
XX 04-NOV-1998; 98US-0107078P.
XX
XX (GEST) GENSET.
XX
XX Griffais R;
XX
XX WPI; 1999-357842/30.
XX
XX
XX Genome sequence of Chlamydia pneumoniae.
PT
XX
XX Page 1131-1132; Disclosure; 1912pp; English.
XX
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
XX frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
XX pneumoniae causes respiratory disease such as pneumonia and bronchitis
XX and is thought to be a contributing factor in heart disease, sarcoidosis,
XX sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
XX polypeptides encoded by the open reading frames of the C. pneumoniae
XX genome (see AAY34584-Y35879) can be used in immunogenic compositions as
XX vaccines. Vectors containing C. pneumoniae nucleotide sequences can also
XX be used as immunogenic compositions, especially where the vector directs
XX the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
XX -OCT-2003 to standardise OS field)
XX
SQ Sequence 282 AA;

Query Match 88.6%; Score 31; DB 2; Length 282;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 7
|:|:|
Db 75 EVNVEETD 81

RESULT 8
ABU27126
ID ABU27126 standard; protein; 282 AA.
XX
XX ABU27126;
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #12653.
DE
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
KW
XX
XX Chlamydia trachomatis.
OS
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002MO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX

PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA30996.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS
 PS Claim 25; SEQ ID NO 55050; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway of
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC parent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 282 AA;
 XX
 Query Match 88.6%; Score 31; DB 6; Length 282;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINLETD 7
 Db 75 EVANVETD 81
 XX
 RESULT 9
 ABUS26773 ID ABUS26773 standard; protein; 282 AA.
 XX
 AC ABUS26773;
 XX
 DT 23-OCT-2003 (revised)
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by prokaryotic essential gene #12300.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Chlamydomphila pneumoniae.

XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 PP
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-0094893P.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA30643.
 DR
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PS
 PS Claim 25; SEQ ID NO 54697; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC parent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 CC
 SQ Sequence 282 AA;
 XX
 Query Match 88.6%; Score 31; DB 6; Length 282;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINLETD 7
 Db 75 EVANVETD 81
 XX
 RESULT 10
 AAY37496

FT /label= OTHER
 FT /note= "OTHER- All Xaa's in this sequence are unknown
 amino acids or the site of a stop codon within the DNA
 sequence"

MO200216439-A2.

28-FEB-2002.

05-MAR-2001; 2001WO-US004941.

07-MAR-2000; 2000US-00519705.
 19-MAY-2000; 2000US-00574454.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Dymnac RT;

WPI; 2002-280918/32.

PT Isolated polynucleotide encoding bone marrow derived polypeptides useful
 for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 disease, and inflammatory bowel disease.

Claim 20; SEQ ID NO 6909; 504bp; English.

This invention relates to a novel isolated polynucleotide comprising a
 nucleotide sequence selected from one of 1680 sequences, a mature protein
 coding portion of them, an active domain of them and their complementary
 sequences. The invention may be useful for the production of compounds
 with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
 immunosuppressive, cytostatic, antiproliferative, antiinflammatory,
 antibacterial, antiviral, antifungal or antiparasitic activity. In
 addition, the disclosed sequences may be useful for gene therapy. The
 polypeptides or their antibodies are useful for treating many diseases
 such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 psoriasis, inflammatory bowel disease and infections caused by bacteria,
 viruses, fungi or parasites. The present sequence is that of a human
 polypeptide of the invention.

Sequence 146 AA;

Query Match 85.7%; Score 30; DB 5; Length 146;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 INLETD 7
 |||||

Db 140 EINLEAD 146

RESULT 15

ADP08273
 ID ADP08273 standard; protein; 272 AA.

AC ADP08273;

DT 26-AUG-2004 (first entry)

DE Neisseria meningitidis MCS8 OMV-related membrane protein - SEQ ID 106.

KM outer-membrane vesicle; antibacterial; antiinflammatory;
 KM meningococcal protein trafficking; localisation; infection; vaccine;
 XX gene therapy.

OS Neisseria meningitidis MCS8.

PN WO2004046177-A2.

PD 03-JUN-2004.

PF 17-NOV-2003; 2003WO-IB006281.

PR 15-NOV-2002; 2002GB-00026734.
 PR 27-MAR-2003; 2003GB-00007131.

PA (CHIR) CHIRON SRL.

PI Norais N, Grandi G;

DR WPI; 2004-420615/39.

PT New compositions having outer-membrane vesicles and proteins from
 PT Neisseria meningitidis, useful in the field of meningococcal
 PT biochemistry, in particular for preventing and/or treating meningococcal
 infections.

Claim 9; SEQ ID NO 106; 79pp; English.

The invention relates to a novel composition comprising outer-membrane
 vesicles (OMV) prepared from a first strain of Neisseria meningitidis and
 1 or more proteins which are present in OMVs prepared from a second
 strain of N. meningitidis, but which are not present in OMVs prepared
 from the first strain. The composition of the invention demonstrates
 antibacterial and antiinflammatory activities and may be useful in the
 field of meningococcal biochemistry, in particular the trafficking and
 localisation of meningococcal proteins, as well as in the prevention and
 treatment of meningococcal infections, possibly via the production of a
 vaccine or gene therapy. The current sequence is that of a Neisseria
 meningitidis MCS8 outer-membrane vesicle (OMV)-related membrane protein
 of the invention.

Sequence 272 AA;

Query Match 85.7%; Score 30; DB 8; Length 272;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 INLETD 7
 |||||

Db 83 INLETD 88

Search completed: August 26, 2005, 01:06:59
 Job time : 334 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2005, 00:56:47 ; Search time 25 Seconds
(without alignments)
26.941 Million cell updates/sec

Title: US-10-726-967a-57

Perfect score: 35

Sequence: 1 EINHETD 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.79:.*
2: p1r1:.*
3: p1r2:.*
4: p1r3:.*
5: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	88.6	282	2 F86577	elongation factor
2	31	88.6	282	2 A72045	translation elonga
3	31	88.6	282	2 A61747	translation elonga
4	31	88.6	282	2 F71484	probable translati
5	31	88.6	390	2 C90288	hypothetical prote
6	31	88.6	928	2 A67749	protein F55C7.7c [
7	31	88.6	2488	2 T42739	guanine nucleotide
8	30	85.7	244	2 T41550	hypothetical prote
9	30	85.7	250	2 S43104	outer membrane pro
10	30	85.7	269	2 A44611	outer membrane pro
11	30	85.7	272	2 C81893	outer membrane pro
12	30	85.7	272	2 A81136	class 5 outer memb
13	30	85.7	403	2 B69196	conserved hypotet
14	30	85.7	437	2 B81844	trigger factor NMA
15	30	85.7	437	2 G81098	trigger factor NMB
16	30	85.7	507	2 D82508	hypothetical prote
17	30	85.7	767	2 A46361	p element homolog
18	30	85.7	872	2 T18651	probable peptide-a
19	29	82.9	35	2 A82151	hypothetical prote
20	29	82.9	74	2 D81796	exodeoxyribonuclea
21	29	82.9	74	2 G81217	exodeoxyribonuclea
22	29	82.9	176	2 F71958	hypothetical prote
23	29	82.9	176	2 B64549	conserved hypotet
24	29	82.9	199	2 S39514	translation elonga
25	29	82.9	218	2 S75585	translation elonga
26	29	82.9	220	2 S73169	translation elonga
27	29	82.9	220	2 H42504	M21 protein - vacc
28	29	82.9	220	2 T28457	hypothetical prote
29	29	82.9	220	2 I36858	O2L protein - vari

30	29	82.9	260	2 G70568	hypothetical prote
31	29	82.9	284	2 G81006	translation elonga
32	29	82.9	287	2 AG2067	hypothetical prote
33	29	82.9	293	2 D89899	elongation factor
34	29	82.9	294	2 I84606	translation elonga
35	29	82.9	294	2 A11281	translation elonga
36	29	82.9	294	2 A61653	translation elonga
37	29	82.9	299	2 B82539	translation elonga
38	29	82.9	306	2 F97120	translation elonga
39	29	82.9	312	2 E87487	translation elonga
40	29	82.9	325	2 T19451	hypothetical prote
41	29	82.9	338	2 I45941	translation elonga
42	29	82.9	346	2 E95258	translation elonga
43	29	82.9	346	2 H98123	elongation factor
44	29	82.9	355	2 C64714	translation elonga
45	29	82.9	355	2 G71804	translation elonga

ALIGNMENTS

RESULT 1

F86577 elongation factor TS [imported] - Chlamydomophila pneumoniae (strain J138)

C/Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: F86577

R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I.

Nucleic Acids Res. 28, 2311-2314, 2000

A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A/Reference number: A86491; MUID:20330349; PMID:10871362

A/Accession: F86577

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-282 <STO>

A/Cross-References: UNIPROT:Q927K8; GB:BA000008; NID:98979065; PIDN:BA98904.1; GSPDB:G

A/Experimental source: strain J138

C/Genetics:

A/Gene: tsf

C/Superfamily: translation elongation factor EF-Ts

Query Match 88.6%; Score 31; DB 2; Length 282;
Best Local Similarity 71.4%; Pred. NO. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EINHETD 7
Db 75 EVNVEITD 81

RESULT 2

A72045 translation elongation factor Ts CP0049 [imported] - Chlamydomophila pneumoniae (strains

N/Alternate names: translation elongation factor EF-Ts

C/Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C/Accession: A72045; D81619

R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.

Nature Genet. 21, 385-389, 1999

A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A/Reference number: A72000; MUID:99206606; PMID:10192388

A/Accession: A72045

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-282 <ARN>

A/Cross-References: UNIPROT:Q927K8; GB:AE001652; GB:AE001363; NID:94376997; PIDN:AA0188

A/Experimental source: strain CML029

R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

; C.; Dodson, R.; Gwim, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg

Nucleic Acids Res. 28, 1397-1406, 2000

A/Title: Genome sequences of Chlamydia trachomatis Morn and Chlamydia pneumoniae AR39.

A/Reference number: A81500; MUID:20150255; PMID:10684935

A/Accession: D81619

A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-282 <REA>
 A:Cross-references: GB:AE002168; GB:AE002161; NID:G7188982; PIDN:AAF37942.1; PID:G718898
 A:Experimental source: strain AK39, HL cells
 C:Genetics:
 A:Gene: tsf; CP0049
 C:Superfamily: translation elongation factor EF-Ts

Query Match 88.6%; Score 31; DB 2; Length 282;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINFETD 7
 Db 75 EVNVEDT 81

RESULT 3
 A81747
 translation elongation factor Ts TC0050 [imported] - Chlamydia muridarum (strain N199)
 C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: A81747
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: A81747
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-282 <TET>
 A:Cross-references: UNIPROT:P71146; GB:AE002272; GB:AE002160; NID:G7190083; PIDN:AAF3893
 C:Genetics:
 A:Gene: TC0050
 C:Superfamily: translation elongation factor EF-Ts

Query Match 88.6%; Score 31; DB 2; Length 282;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINFETD 7
 Db 75 EVNVEDT 81

RESULT 4
 F71484
 Probable translation elongation factor EF-Ts - Chlamydia trachomatis (serotype D, strain C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
 C:Accession: F71484
 R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.
 A:Reference number: A71570; MUID:9900809; PMID:9784336
 A:Accession: F71484
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-282 <ARN>
 A:Cross-references: UNIPROT:O84666; GB:AE001338; GB:AE001273; NID:G3329126; PIDN:AA6827
 A:Experimental source: serotype D, strain UW-3/Cx
 C:Genetics:
 A:Gene: tsf
 C:Superfamily: translation elongation factor EF-Ts

Query Match 88.6%; Score 31; DB 2; Length 282;
 Best Local Similarity 71.4%; Pred. No. 26;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINFETD 7

Db 75 EVNVEDT 81

RESULT 5
 C90288
 hypochetrical protein SSO1325 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C:Accession: C90288
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-arrett, R.A.; Ragan, M.A.; Jensen, C.W.; Van der Oost, J.
 Submitted to Genbank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: C90288
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-390 <KUR>
 A:Cross-references: UNIPROT:Q97YU7; GB:AE006641; NID:G13814532; PIDN:AAK41562.1; GSPDB: C:Genetics:
 A:Gene: SSO1325

Query Match 88.6%; Score 31; DB 2; Length 390;
 Best Local Similarity 85.7%; Pred. No. 38;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINFETD 7
 Db 191 EINFETD 197

RESULT 6
 A87749
 protein F55C7.7c [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C:Accession: A87749
 R:Anonymous. The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: A87749
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-928 <STO>
 A:Cross-references: UNIPROT:O61528; GB:chr I; PIDN:AACT1110.1; PID:G3844604; GSPDB:GN00C
 C:Genetics:
 A:Gene: F55C7.7c
 A:Map position: 1

Query Match 88.6%; Score 31; DB 2; Length 928;
 Best Local Similarity 71.4%; Pred. No. 1e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EINFETD 7
 Db 651 QVNFETD 657

RESULT 7
 T42739
 guanine nucleotide exchange factor unc-73a - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T42739; T42740; T34427; T34430; T34426
 R:Steven, R.; Kubisek, T.J.; Zheng, H.; Kulkarni, S.; Mancillas, J.; Ruiz Morales, A.; Cell 92, 785-795, 1998
 A:Title: UNC-73 activates the Rac GTPase and is required for cell and growth cone migrat

A:Reference number: 222556; MUID:98188103; PMID:9529254
A:Accession: T42739
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2488 <STE>
A:Cross-references: UNIPROT:O61528; EMBL:AF048834; NID:G2944395; PIDN:AAIC12931.1; PID:G2944395
A:Accession: T42740
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1634, 'ON', '1637', 'E' <ST2>
A:Cross-references: EMBL:AF048835; NID:G2944397; PIDN:AAIC12932.1; PID:G2944398
R:Du, Z.; Le, T.
Submitted to the EMBL Data Library, November 1996
A:Description: The sequence of *C. elegans* cosmid F55C7.
A:Reference number: 221524
A:Accession: T34427
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2488 <DUZ>
A:Cross-references: EMBL:U80436; PIDN:AACT1108.1; GSPDB:GN00019; CESP:F55C7.7a
A:Experimental source: strain Bristol N2; clone F55C7
A:Accession: T34430
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1634, 'ON', '1637', 'E' <DU2>
A:Cross-references: EMBL:U80436; PIDN:AACT1109.1; GSPDB:GN00019; CESP:F55C7.7b
A:Experimental source: strain Bristol N2; clone F55C7
A:Accession: T34426
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 'MWIKCF', '1568', 'LRK', '1572-2488' <DU3>
A:Cross-references: EMBL:U80436; PIDN:AACT1110.1; GSPDB:GN00019; CESP:F55C7.7c
A:Experimental source: strain Bristol N2; clone F55C7
C:Genetics:
A:Gene: unc-73; F55C7.7a; unc-73; F55C7.7b; CESP:F55C7.7c
A:Map position: 1
A:Intons: 7/3; 64/2; 95/3; 230/3; 265/2; 335/3; 360/3; 488/3; 569/1; 695/1; 1028/2; 10619/3; 2262/2; 2331/1; 2351/3; 2396/1; 2431/1
Query Match 88.6%; Score 31; DB 2; Length 2488;
Best Local Similarity 71.4%; Pred. No. 3.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
CY 1 EINFETD 7
Db 2211 QVNFETD 2217

RESULT 8
T41550
hypothetical protein SPCC70.04c - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41550
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.
Submitted to the EMBL Data Library, June 1998
A:Reference number: 222001
A:Accession: T41550
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-244 <WOO>
A:Cross-references: UNIPROT:O74525; EMBL:AL023794; PIDN:CAA19354.1; GSPDB:GN00068; SPDB:
C:Experimental source: strain 972h-; cosmid c70
C:Genetics:
A:Gene: SPDB:SPCC70.04c
A:Map position: 3
A:Intons: 146/1; 202/1
C:Superfamily: *Schizosaccharomyces pombe* hypothetical protein SPCC70.04c
Query Match 85.7%; Score 30; DB 2; Length 244;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINFETD 7
Db 15 EINFETD 21

RESULT 9
S43104
outer membrane protein class 5c - *Neisseria meningitidis* (fragment)
C:Species: *Neisseria meningitidis*
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S43104
R:Guillen, G.; Leal, M.J.; Alvarez, A.; Delgado, M.; Silva, R.; Herrera, L.
Submitted to the EMBL Data Library, March 1994
A:Description: Cloning and expression of the *Neisseria meningitidis* 5C outer membrane p.
A:Reference number: S43104
A:Accession: S43104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-250 <GUI>
A:Cross-references: UNIPROT:Q51228; EMBL:X78221; NID:G467599; PIDN:CAA55065.1; PID:G984

Query Match 85.7%; Score 30; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 2 EINFETD 7
Db 64 EINFETD 69

RESULT 10
A44611
outer membrane protein class 5c precursor - *Neisseria meningitidis* (strain Z3476)
C:Species: *Neisseria meningitidis*
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A44611
R:Olyhoek, A.J.M.; Sarkari, J.; Bopp, M.; Morelli, G.; Achtmann, M.
Microb. Pathog. 11, 249-257, 1991
A:Title: Cloning and expression in *Escherichia coli* of *opc*, the gene for an unusual class
A:Reference number: A44611; MUID:92261288; PMID:1813777
A:Accession: A44611
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-269 <OLY>
A:Cross-references: UNIPROT:Q51227; GB:M80195
C:Genetics:
A:Gene: *opc*

Query Match 85.7%; Score 30; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 2 EINFETD 7
Db 83 EINFETD 88

RESULT 11
C81893
outer membrane protein precursor NMA1251 [imported] - *Neisseria meningitidis* (strain Z2)
C:Species: *Neisseria meningitidis*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: C81893
R:Parkhill, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More
; Holroyd, S.; Oegels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A:Reference number: A8175; MUID:20222556; PMID:10761919
A:Accession: C81893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <PAR>
A:Cross-references: UNIPROT:Q51227; GB:AL162755; GB:AL157959; NID:G7379742; PIDN:CAB845C

A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: opca; NMA1251

Query Match
Best Local Similarity 85.7%; Score 30; DB 2; Length 272;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 INLETD 7
Db 83 INLETD 88

RESULT 12

A:1126
Class 5 Outer membrane protein NMB1053 [imported] - Neisseria meningitidis (strain MCS8)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: A81126
R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Scienc, 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Va
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: A81126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <TET>
A:Cross-references: UNIPROT:Q51229; GB:AE002456; GB:AE002098; NID:g7226293; PIDN:AAF4145
A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1053

Query Match
Best Local Similarity 85.7%; Score 30; DB 2; Length 272;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 INLETD 7
Db 83 INLETD 88

RESULT 13

BC9196
conserved hypothetical protein MTH72 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C:Accession: B69196
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Kl, S.; Church, G.W.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: B69196
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-403 <MTH>
A:Cross-references: UNIPROT:Q26176; GB:AE000798; GB:AE000666; NID:g2621094; PIDN:AA8457
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH72

C:Superfamily: tetratricopeptide repeat homology
F:16-49/Domain: tetratricopeptide repeat homology <TT01>
F:50-83/Domain: tetratricopeptide repeat homology <TT02>
F:84-117/Domain: tetratricopeptide repeat homology <TT03>
F:118-151/Domain: tetratricopeptide repeat homology <TT04>
F:152-185/Domain: tetratricopeptide repeat homology <TT05>
F:186-219/Domain: tetratricopeptide repeat homology <TT06>
F:220-253/Domain: tetratricopeptide repeat homology <TT07>
F:254-287/Domain: tetratricopeptide repeat homology <TT08>

F:288-321/Domain: tetratricopeptide repeat homology <TT09>
F:322-355/Domain: tetratricopeptide repeat homology <TT10>
F:356-389/Domain: tetratricopeptide repeat homology <TT11>

Query Match
Best Local Similarity 85.7%; Score 30; DB 2; Length 403;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EINLETD 7
Db 282 EINLETD 288

RESULT 14

B81844
trigger factor NMA1526 [imported] - Neisseria meningitidis (strain Z2491 serogroup A)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: B81844
R:Parthill, U.; Achtman, M.; James, K.D.; Bentley, S.D.; Church, C.; Klee, S.R.; Morel
; Holroyd, S.; Jørgens, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: B81844
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <PAR>
A:Cross-references: UNIPROT:Q9JU32; GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CAB8475
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: tfg; NMA1526

F:163-207/Domain: BKB-type peptidyl-prolyl cis-trans isomerase (trigger factor); BKB-type

Query Match
Best Local Similarity 85.7%; Score 30; DB 2; Length 437;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EINLETD 7
Db 24 EINLETD 30

RESULT 15

G81098
trigger factor NMB1313 [imported] - Neisseria meningitidis (strain MCS8 serogroup B)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: G81098
R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Scienc, 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Va
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: G81098
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-437 <TET>
A:Cross-references: UNIPROT:Q9J237; GB:AE002479; GB:AE002098; NID:g7226543; PIDN:AAF4166
A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1313

C:Superfamily: FKBP-type peptidyl-prolyl cis-trans isomerase (trigger factor); BKB-type
F:163-207/Domain: BKB-type peptidyl-prolyl isomerase homology <PPI>

Query Match
Best Local Similarity 85.7%; Score 30; DB 2; Length 437;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EINLETD 7

Db ||| |||
 24 EINFELD 30

Search completed: August 26, 2005, 01:09:15
Job time : 26 secs

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2005, 00:30:41 ; Search time 98 Seconds
(without alignments)
36.577 Million cell updates/sec

Title: US-10-726-967a-57
Perfect score: 35
Sequence: 1 EINLETD 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	954	1 GOSP_SYNEL	Q6d113 synechococ
2	31	88.6	202	2 06QIHI	Q6q1h1 chlamydia t
3	31	88.6	202	2 06QIHI2	Q6q1h2 chlamydia t
4	31	88.6	241	2 08G6H7	Q8g6h7 bifidobacte
5	31	88.6	282	1 EPTS_CHICV	Q824u4 chlamydophi
6	31	88.6	282	1 EPTS_CHLMU	P71146 chlamydia m
7	31	88.6	282	1 EPTS_CHLPN	Q927f8 chlamydia p
8	31	88.6	282	1 EPTS_CHLTP	Q846b6 chlamydia t
9	31	88.6	390	2 097YV7	Q97YV7 sulfobolus
10	31	88.6	842	2 06C199	Q6c199 yarrowia li
11	31	88.6	928	2 07JNG6	Q7jng6 caenorhabdi
12	31	88.6	2488	2 061528	Q61528 caenorhabdi
13	30	85.7	98	2 067JBO	Q67jbo symbiobacte
14	30	85.7	151	2 06A1S9	Q6a1s9 pseudomonas
15	30	85.7	151	2 06A1T2	Q6a1t2 pseudomonas
16	30	85.7	151	2 06A1T6	Q6a1t6 pseudomonas
17	30	85.7	151	2 06A1U4	Q6a1u4 pseudomonas
18	30	85.7	151	2 06A1V0	Q6a1v0 pseudomonas
19	30	85.7	151	2 06A1V3	Q6a1v3 pseudomonas
20	30	85.7	156	2 099ME7	Q99me7 rattus norv
21	30	85.7	187	2 07T9Y9	Q7t9y9 adoxophyes
22	30	85.7	188	2 06RSY7	Q6rsy7 uncultured
23	30	85.7	189	2 089Y77	Q89y77 uncultured
24	30	85.7	201	2 082S10	Q82s10 nitrosomonas
25	30	85.7	202	1 EPTS_PHATR	Q9tk50 phaeodactyl
26	30	85.7	208	1 EPTS_CYACA	Q9cm30 cyanidium c
27	30	85.7	219	1 EPTS_SYNPX	Q7u734 synechococ
28	30	85.7	220	2 093T92	Q93t92 alcaligenes
29	30	85.7	220	2 06VUY6	Q6vuy6 uncultured
30	30	85.7	220	2 06VV01	Q6vv01 uncultured
31	30	85.7	220	2 09AP72	Q9ap72 uncultured

32	30	85.7	232	2	Q93S43	Q93s43 neisseria m
33	30	85.7	235	2	Q93S45	Q93s45 neisseria m
34	30	85.7	238	2	Q93S44	Q93s44 neisseria m
35	30	85.7	240	2	Q7UIR3	Q7ui3 rhodospirill
36	30	85.7	244	2	Q74525	Q74525 schizosacch
37	30	85.7	248	2	Q93S42	Q93s42 neisseria m
38	30	85.7	250	2	Q51228	Q51228 neisseria m
39	30	85.7	252	2	Q9AE79	Q9ae79 neisseria m
40	30	85.7	259	1	EPTS_CANBF	Q7vres candidatus
41	30	85.7	272	2	Q51230	Q51230 neisseria m
42	30	85.7	272	2	Q6O8T8	Q6o8t8 uncultured
43	30	85.7	272	2	Q51227	Q51227 neisseria m
44	30	85.7	272	2	Q51229	Q51229 neisseria m
45	30	85.7	272	2	Q7AR44	Q7ar44 neisseria m

ALIGNMENTS

```

RESULT 1
GOSP_SYNEL          STANDARD;          PRT;          954 AA.
ID      GOSP_SYNEL
AC      Q6D113;
DT      10-OCT-2003 (Rel. 42, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Glycine dehydrogenase [decarboxylating] (EC 1.4.4.2) (Glycine
DE      decarboxylase) (Glycine cleavage system P-protein).
GN      Namegcvp; OrderedLocustNames=tl1603;
OS      Synechococcus elongatus (Thermosynechococcus elongatus).
OC      Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX      NCBI_TaxID=32046;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BP-1; Pubmed=12240834;
RX      MEDLINE=2225144;
RA      Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA      Watanabe A., Iritaguchi M., Kawashima K., Kimura T., Kishida Y.,
RA      Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA      Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT      "complete genome structure of the thermophilic cyanobacterium
RT      Thermosynechococcus elongatus BP-1."
RL      DNA Ref. 9:123-130(2002).
CC      -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC      glycine. The P protein binds the alpha-amino group of glycine
CC      through its pyridoxal phosphate cofactor; CO(2) is released and
CC      the remaining methylamine moiety is then transferred to the
CC      liposamide cofactor of the H protein (By similarity).
CC      -1- CATALYTIC ACTIVITY: Glycine + H-protein-lipoalysine = H-protein-
CC      S-aminomethylidihydroalipoyllysine + CO(2).
CC      -1- COFACTOR: Pyridoxal phosphate (By similarity).
CC      -1- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC      P, T, L and H (By similarity).
CC      -1- SIMILARITY: Belongs to the gcvp family.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; AP005374; BAC09155.1; -.
DR      HAMAP; MF_00711; -; 1.
DR      InterPro; IPR003437; GDC-P.
DR      InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR      Pfam; PF02347; GDC-P; 1.
DR      TIGRfam; TIGR00461; gcvp; 1.
KW      Complete proteome; Oxidoreductase; Pyridoxal phosphate.
FT      BINDING 706 Pyridoxal phosphate (By similarity).
SQ      SEQUENCE 954 AA; 103794 MW; 01AE250CB6ACF95 CRC64;

```

Query Match 100.0%; Score 35; DB 1; Length 954;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINHETD 7
 DB 10 EINHETD 16

RESULT 2

06QIH1 PRELIMINARY; PRT; 202 AA.
 AC 06QIH1; PRT; 202 AA.
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Elongation factor TS (Fragment).
 GN Name=flin; ORFNames=CT679;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxId=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R/Har-1;
 RA Brunelle B.W., Nicholson T.L., Stephens R.S.;
 CC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the
 exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-
 Tu.GTP complex up to the GTP hydrolysis stage on the ribosome (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the EF-Ts family.
 DR EMBL; AY539791; AAS90254.1; -
 DR GO; GO:0003746; P:translation elongation factor activity; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR001816; P:translational elongation; IEA.
 DR Pfam; PF00889; EF-TS; 1-TS.
 DR PROSITE; PS01127; EF-TS_2; 1.
 DR Elongation factor; Protein biosynthesis.
 KW NON_TER 1
 FT SEQUENCE 202 AA; 21941 MW; FD16EAB3A4EC097D CRC64;
 SQ

Query Match 88.6%; Score 31; DB 2; Length 202;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINHETD 7
 DB 12 EVNVEDT 18

RESULT 3

06QIH2 PRELIMINARY; PRT; 202 AA.
 AC 06QIH2; PRT; 202 AA.
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Elongation factor TS (Fragment).
 GN Name=flin; ORFNames=CT679;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxId=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R/Har-1;
 RA Brunelle B.W., Nicholson T.L., Stephens R.S.;
 CC Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the
 exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-
 Tu.GTP complex up to the GTP hydrolysis stage on the ribosome (By

CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the EF-Ts family.
 DR EMBL; AY539790; AAS90254.1; -
 DR GO; GO:0003746; P:translation elongation factor activity; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR001816; P:translational elongation; IEA.
 DR Pfam; PF00889; EF-TS; 1-TS.
 DR PROSITE; PS01127; EF-TS_2; 1.
 DR Elongation factor; Protein biosynthesis.
 KW NON_TER 1
 FT SEQUENCE 202 AA; 21925 MW; FD16EAB3A3764D1 CRC64;
 SQ

Query Match 88.6%; Score 31; DB 2; Length 202;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINHETD 7
 DB 12 EVNVEDT 18

RESULT 4

08G6H7 PRELIMINARY; PRT; 241 AA.
 AC 08G6H7; PRT; 241 AA.
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DE Possible voltage-gated potassium channel protein.
 GN OrderedLocustNames=Bl00664;
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxId=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCC 2705;
 RA MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
 RA Schell M.A., Karamiantzou M., Snel B., Vilanova D., Berger B.,
 RA Pessl G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
 RA Pridmore R.D., Arigoni F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation
 to the human gastrointestinal tract."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 DR EMBL; AE014688; AAN24485.1; -
 DR HSSP; Q54397; 1095.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008076; C:voltage-gated potassium channel complex; IEA.
 DR GO; GO:0005249; F:voltage-gated potassium channel activity; IEA.
 DR GO; GO:0006812; P:cation transport; IEA.
 DR GO; GO:0006813; P:potassium ion transport; IEA.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001622; K+channel_pore.
 DR InterPro; IPR003091; K+channel.
 DR InterPro; IPR005820; M+channel_nlg.
 DR Pfam; PF00520; Ion_trans_1.
 DR PRINTS; PRO0169; KCHANNEL.
 KW Complete proteome; Ion transport; Ionic channel; Transmembrane;
 KW TRANSPORT.
 SQ SEQUENCE 241 AA; 26902 MW; FBS940B31P24A712 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 241;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINHETD 7
 DB 184 QVNLFTD 190

```

RESULT 5
EFTS_CHLCV STANDARD: PRT: 282 AA.
ID EFTS_CHLCV STANDARD: PRT: 282 AA.
AC Q824U4:
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Elongation factor Ts (EF-Ts).
GN Name:tsf; OrderedLocustNames=CCA00045;
OS Chlamydia caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzaple E.K., Knouri H.M., Federova N.B.,
RA Carty H.A., Umeyam L.A., Haft D.H., Peterson J.D., Beaman M.J.,
RA White O., Salzberg S.L., Heia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydia caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the
CC exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-
CC Tu.GTP complex up to the GTP hydrolysis stage on the ribosome.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the EF-Ts family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016994; AAP04797.1; -.
DR HSSP; P02997; IEFU.
DR TIGR; CCA00045; -.
DR HAMAP; MF_00050; -.
DR InterPro; IPR001816; EF-TS.
DR InterPro; IPR000449; UBA.
DR InterPro; IPR009060; UBA_like.
DR Pfam; PF00889; EF-TS; 1.
DR Pfam; PF00627; UBA; 1.
DR TIGRPFAMs; TIGR00116; tsf; 1.
DR PROSITE; PS01126; EF-TS_1; 1.
DR PROSITE; PS01127; EF-TS_2; 1.
DR Complete proteome; Elongation factor; Protein biosynthesis.
KM Complete proteome; Elongation factor; Protein biosynthesis.
FT SITE 80 Involved in Mg(2+) ion dislocation from
FT SITE 83 EF-Tu (By similarity).
SQ SEQUENCE 282 AA; 30478 MW; 3D09BFBF811A5DA7 CRC64;
Query Match 88.6%; Score 31; DB 1; Length 282;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 EINFETD 7
Db 75 EVNVEITD 81

```

```

GN Name:tsf; OrderedLocustNames=TC0050;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOBn;
RX MEDLINE=97386531; PubMed=9244380; DOI=10.1006/abbi.1997.0178;
RA Zhang Y.X., Tao J., Zhou M., Meng Q., Zhang L., Shen L., Klein R.,
RA Miller D.L.;
RT "Elongation factor Ts of Chlamydia trachomatis: structure of the gene
RT and properties of the protein."
RL Arch. Biochem. Biophys. 344:43-52(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MOBn; Nig9;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Uetback T.R., Berry K.J.,
RA Baes S., Linher K.D., Wetman J.F., Knouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOBn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the
CC exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-
CC Tu.GTP complex up to the GTP hydrolysis stage on the ribosome.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the EF-Ts family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60196; AAB07070.1; -.
DR EMBL; AE002272; AAF38939.1; -.
DR PIR; A81747; A81747.
DR HSSP; P43895; ITFE.
DR TIGR; TC0050; -.
DR HAMAP; MF_00050; -.
DR InterPro; IPR001816; EF-TS.
DR InterPro; IPR000449; UBA.
DR InterPro; IPR009060; UBA_like.
DR Pfam; PF00889; EF-TS; 1.
DR Pfam; PF00627; UBA; 1.
DR TIGRPFAMs; TIGR00116; tsf; 1.
DR PROSITE; PS01126; EF-TS_1; 1.
DR PROSITE; PS01127; EF-TS_2; 1.
DR Complete proteome; Elongation factor; Protein biosynthesis.
KM Complete proteome; Elongation factor; Protein biosynthesis.
FT SITE 80 Involved in Mg(2+) ion dislocation from
FT SITE 83 EF-Tu (By similarity).
SQ SEQUENCE 282 AA; 30824 MW; AE56D44808951B17 CRC64;
Query Match 88.6%; Score 31; DB 1; Length 282;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 EINFETD 7
Db 75 EVNVEITD 81

```

```

RESULT 6
EFTS_CHLMU STANDARD: PRT: 282 AA.
ID EFTS_CHLMU STANDARD: PRT: 282 AA.
AC P71145;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Elongation factor Ts (EF-Ts).

```

```

RESULT 7
EFTS_CHLPN STANDARD: PRT: 282 AA.
ID EFTS_CHLPN STANDARD: PRT: 282 AA.
AC Q927K8; Q9JSC0;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Elongation factor Ts (EF-Ts).
 GN Name: OrderedLocustNames=Cp00697, CP0049, Cp0074;
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_TaxID=83558;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=CML029;
 RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
 RA Kalman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
 RT Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RL "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."; Nat. Genet. 21:385-389(1999).
 RN (2) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RT White O., Hickey E.K., Peterson J.D., Uppertack T.R., Berry K.J.,
 RA Bass S., Linher K.D., Weidman J.F., Khouiri H.M., Craven B., Bowman C.,
 RT Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
 RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39."; Nucleic Acids Res. 28:1397-1406(2000).
 RN (3) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RT Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
 RL "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CML029 from USA."; Nucleic Acids Res. 28:2311-2314(2000).
 RN (4) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=TW-183;
 RA Geng M.M., Schumacher A., Muehlbauer I., Bensch K.W., Schaefer K.P.,
 RT Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
 RL "The genome sequence of Chlamydia pneumoniae TW183 and comparison with other Chlamydia strains based on whole genome sequence analysis."; Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-Tu.GTP complex up to the GTP hydrolysis stage on the ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the EF-Ts family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001652; AAD18836.1; -;
 DR EMBL; AE002167; AAF37942.1; -;
 DR EMBL; AP002547; BAA98904.1; -;
 DR EMBL; AE017159; AAP98653.1; -;
 DR PIR; A72045; A72045.
 DR PIR; F86577; F86577.
 DR HSSP; P02997; 1EPU.
 DR PHCT-2DPAGE; Q927K8; -;
 DR TIGR; CP0049; -;
 DR HAMAP; MF_00050; -; 1.
 DR InterPro; IPR001816; EF_TS.
 DR InterPro; IPR000449; UBA.
 DR InterPro; IPR009060; UBA_like.
 DR Pfam; PF00889; EF_TS; 1.
 DR Pfam; PF00627; UBA; 1.
 DR TIGRPFAM; TIGR00116; tsf; 1.

DR PROSITE; PS01126; EF_TS_1; 1.
 DR PROSITE; PS01127; EF_TS_2; 1.
 KW Complete proteome; Elongation factor; Protein biosynthesis.
 FT SITE 80 Involved in Mg(2+) ion displacement from EF-Tu (By similarity).
 FT CONFLICT 71 T -> A (in Ref. 3).
 SO SEQUENCE 282 AA; 30392 MW; 1609C93A34AD5979 CRC64;
 Query Match 88.6%; Score 31; DB 1; Length 282;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 EINFELD 7
 Db 75 EVANVED 81
 RESULT 8
 ID EFTS_CHLTR STANDARD; PRT; 282 AA.
 AC 084686;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Elongation factor Ts (EF-Ts).
 GN Name: OrderedLocustNames=CT679;
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 ON NCBI_TaxID=813;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=D/TW-3/Cx;
 RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA Davis R.W.;
 RL "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis."; Science 282:754-759(1998).
 CC -1- FUNCTION: Associates with the EF-Tu.GDP complex and induces the exchange of GDP to GTP. It remains bound to the aminoacyl-tRNA.EF-Tu.GTP complex up to the GTP hydrolysis stage on the ribosome.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the EF-Ts family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001338; AAC68274.1; -;
 DR PIR; F71484; F71484.
 DR HSSP; P43895; 1TPE.
 DR PHCT-2DPAGE; O84686; -;
 DR HAMAP; MF_00050; -; 1.
 DR InterPro; IPR001816; EF_TS.
 DR InterPro; IPR000449; UBA.
 DR InterPro; IPR009060; UBA_like.
 DR Pfam; PF00889; EF_TS; 1.
 DR Pfam; PF00627; UBA; 1.
 DR TIGRPFAM; TIGR00116; tsf; 1.
 DR PROSITE; PS01126; EF_TS_1; 1.
 DR PROSITE; PS01127; EF_TS_2; 1.
 KW Complete proteome; Elongation factor; Protein biosynthesis.
 FT SITE 80 Involved in Mg(2+) ion displacement from EF-Tu (By similarity).
 FT CONFLICT 83 T -> A (in Ref. 3).
 SO SEQUENCE 282 AA; 30882 MW; 882A5AF38F0D1F02 CRC64;
 Query Match 88.6%; Score 31; DB 1; Length 282;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EINFETD 7
|:|:|
Db 75 EVNVEDT 81

RESULT 9

O97YJ7 PRELIMINARY; PRT; 390 AA.
AC O97YJ7;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SSO1325;
OS Sulfolobus solfataricus.
OC Archaeae; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
CC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AB006747; AAK41562.1; -.
DR PIR; C90288; C90288.
SQ Complete proteome; Hypothetical protein.
SC SEQUENCE 390 AA; 45736 MW; 925C3071091C98A3 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 390;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EINFETD 7
|:|:|
Db 191 EINFETD 197

RESULT 10

O6C199 PRELIMINARY; PRT; 842 AA.
AC O6C199;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Yarrowia lipolytica chromosome A of strain CL1899 of Yarrowia
lipolytica.
GN ORFNames=YALI0A003529;
OS Yarrowia lipolytica CL1899.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Dipodasaccaeae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL1899;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne M., Blyksten C.,
RA Boissiere A., Boyer J., Cattolico L., Confalonieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hartraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Kozul R., Lemaire M., Lesur I., Ma L., Muller H.,

RA Nicaud J.M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Weschoff E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=CL1899;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382127; CAG83532.1; -.
DR GO; GO:0005525; F:GTP binding; IEA.
DR GO; GO:0003746; F:translation elongation factor activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006414; P:translational elongation; IEA.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProTSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C_1.
DR Pfam; PF03764; EFG_IV_1.
DR Pfam; PF00009; GTP_BFTU_1.
DR Pfam; PF03144; GTP_BFTU_D2_1.
DR PRINTS; PRO0315; EIDONGATNCT.
DR TIGRFAMs; TIGR00231; small_GTP_1.
DR PROSITE; PS00301; EFATOR_GTP_1.
DR GTP-binding; Protein biosynthesis.
SQ SEQUENCE 842 AA; 93402 MW; 36FE46655FD3453 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 842;
Best Local Similarity 85.7%; Pred. No. 6.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EINFETD 7
|:|:|
Db 304 EINFETD 310

RESULT 11

O7JNG6 PRELIMINARY; PRT; 928 AA.
AC O7JNG6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Uncoordinated protein 73, isoform c.
GN Name=unc-73; ORFNames=F55C7.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
CC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z., Le T.;
RT "The sequence of C. elegans cosmid F55C7.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.

```

RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80436; AAC71110.1; -.
DR WormBase; WBGene0006805; F55C7.7.
DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR003961; FN_III
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhogEF.
DR InterPro; IPR01511; SH3_2.
DR Pfam; PF00041; fn3_1.
DR Pfam; PF00621; RhogEF; 1.
DR Pfam; PF07653; SH3_2; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhogEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00741; DH_1; 1.
DR PROSITE; PS0010; DH_2; 1.
DR PROSITE; PS50853; FN3; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 928 AA; 102434 MW; 62E6F4673F74D6E CRC64;

Query Match 88.6%; Score 31; DB 2; Length 928;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINFLETD 7
Db 651 QVNFLETD 657

RESULT 12
O61528 PRELIMINARY; PRT; 2488 AA.
AC O61528; O61529; P91339.
DT 01-AUG-1998 (TREMBL)rel. 07, Created
DT 01-AUG-1998 (TREMBL)rel. 07, Last sequence update
DT 25-OCT-2004 (TREMBL)rel. 28, Last annotation update
DE Guanine nucleotide exchange factor UNC-73A (uncoordinated protein 73,
DE isoform a).
GN Name=unc-73; ORFNames=F55C7.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98188103; PubMed=9529254; DOI=10.1016/S0092-8674(00)81406-3;
RA Steven R., Kubisecki T.J., Zheng H., Kulkarni S., Mancillas J.,
RA Ruiz Morales A., Hogue C.W.V., Pawson T., Culotti J.;
RT "UNC-73 activates the Rac GTPase and is required for cell and growth
RT cone migrations in C. elegans.";
RL Cell 92:785-795(1998).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Du Z., Le T.;
RT "The sequence of C. elegans cosmid F55C7.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AF048834; AAC12931.1; -.
DR EMBL; U80436; AAC71108.1; -.
DR PIR; A87749; A87749.
DR PIR; D87749; D87749.
DR PIR; T34428; T34428.
DR PIR; T42739; T42739.
DR HSP; Q64036; IKZG.
DR WormBase; WBGene0006805; F55C7.
DR WormBep; F55C7.7a; CE19464.
DR GO; GO:0005085; F:guanyl-nucleotide exchange factor activity; IEA.
DR GO; GO:007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR001251; CRAL_TRIO_C.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001331; GDS_CDC24.
DR InterPro; IPR003599; IG_1like.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000219; RhogEF.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR01511; SH3_2.
DR InterPro; IPR02017; Spectrin.
DR Pfam; PF00041; fn3_1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00621; RhogEF; 2.
DR Pfam; PF07653; SH3_2; 1.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00233; PH; 2.
DR SMART; SM00325; RhogEF; 2.
DR SMART; SM00516; SEC14; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00150; SPC2.
DR PROSITE; PS00741; DH_1; 1.
DR PROSITE; PS0010; DH_2; 1.
DR PROSITE; PS50853; FN3; 1.
DR PROSITE; PS50011; GLA_1; UNKNOWN_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 2488 AA; 282876 MW; FEF3F93525465C37 CRC64;

Query Match 88.6%; Score 31; DB 2; Length 2488;
Best Local Similarity 71.4%; Pred. No. 2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```


QY 1 INLETD 7
:|||||
Db 2211 QVLETD 2217

RESULT 13

Q67JEB0 PRELIMINARY; PRT; 98 AA.
AC Q67JEB0; 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN ORFNames=STH3228;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsui T.,
Moriwaka K., Ikeda H., Hattori M., Bepko T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
thermophilum";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
KM EMBL; AP006840; BAD42210.1; -.
SQ SEQUENCE 98 AA; 11145 MW; 9D808777974633FE CRC64;

Query Match 85.7%; Score 30; DB 2; Length 98;
Best Local Similarity 71.4%; Pred. No. 96;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 INLETD 7
:|||||
Db 36 EVNLETD 42

RESULT 14

Q6A1S9 PRELIMINARY; PRT; 151 AA.
AC Q6A1S9; 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Nitrous-oxide reductase (EC 1.7.99.6) (Fragment).
GN Name=nos2;
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RA Cladera A.M., Bannas A., Barcelo M., Lalucat J., Garcia-Valdes E.;
RT "Comparative Genetic Diversity of Pseudomonas stutzeri Genomovars,
RT Clonal Structure, and Phylogeny of the Species.";
RL J. Bacteriol. 186:5239-5248(2004).
EMBL: AJ631996; CAF74911.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0050304; F:nitrous-oxide reductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR011045; N2O_reductase_N.
DR ProDom; PD000131; Copper_CuA; 1.
KM Oxidoreductase.
FT NON_TER 1
FT NON_TER 151
SQ SEQUENCE 151 AA; 17156 MW; 942FBF8CF387FC9B CRC64;

Query Match 85.7%; Score 30; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 INLETD 7
:|||||
Db 56 INLETD 61

RESULT 15

Q6A1T2 PRELIMINARY; PRT; 151 AA.
AC Q6A1T2; 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Nitrous-oxide reductase (EC 1.7.99.6) (Fragment).
GN Name=nos2;
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RA Cladera A.M., Bannas A., Barcelo M., Lalucat J., Garcia-Valdes E.;
RT "Comparative Genetic Diversity of Pseudomonas stutzeri Genomovars,
RT Clonal Structure, and Phylogeny of the Species.";
RL J. Bacteriol. 186:5239-5248(2004).
EMBL: AJ631993; CAF74908.1; -.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0050304; F:nitrous-oxide reductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR011045; N2O_reductase_N.
DR ProDom; PD000131; Copper_CuA; 1.
KM Oxidoreductase.
FT NON_TER 1
FT NON_TER 151
SQ SEQUENCE 151 AA; 17086 MW; 2A3F016686CC3EA2 CRC64;

Query Match 85.7%; Score 30; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 INLETD 7
:|||||
Db 56 INLETD 61

Search completed: August 26, 2005, 01:08:44
Job time : 100 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2005, 00:59:56 ; Search time 29 Seconds
(without alignments)
18.019 Million cell updates/sec

Title: US-10-726-967a-57

Perfect score: 35

Sequence: 1 EINHETD 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_Aa.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	88.6	282	4	US-09-198-452A-739
2	31	88.6	287	4	US-09-438-185A-699
3	30	85.7	146	4	US-09-270-767-36607
4	30	85.7	146	4	US-09-270-767-51824
5	30	85.7	333	4	US-08-991-761A-8
6	30	85.7	964	4	US-09-543-681A-5035
7	29	82.9	201	4	US-09-248-796A-16599
8	29	82.9	208	3	US-09-134-001C-4971
9	29	82.9	210	4	US-09-902-540-15089
10	29	82.9	247	4	US-09-248-796A-22971
11	29	82.9	273	4	US-09-710-279-752
12	29	82.9	307	4	US-09-710-279-2668
13	29	82.9	314	4	US-09-248-796A-18201
14	29	82.9	319	3	US-09-134-001C-5553
15	29	82.9	346	4	US-09-583-110-4370
16	29	82.9	359	4	US-09-107-433-2752
17	29	82.9	382	4	US-09-949-016-10652
18	28	80.0	215	4	US-09-134-000C-4001
19	28	80.0	308	3	US-09-134-001C-2987
20	28	80.0	318	4	US-09-543-681A-6384
21	28	80.0	334	4	US-09-107-532A-7224
22	28	80.0	346	4	US-09-543-681A-4190
23	28	80.0	357	4	US-09-248-796A-15519
24	28	80.0	540	4	US-09-107-532A-5314
25	28	80.0	615	4	US-09-248-796A-14606
26	28	80.0	851	4	US-09-619-353-12
27	28	80.0	864	4	US-09-134-000C-6025

28	28	80.0	885	4	US-09-248-796A-14427	Sequence 14427, A
29	27	77.1	133	4	US-09-513-999C-5226	Sequence 5226, Ap
30	27	77.1	160	4	US-09-270-767-35707	Sequence 35707, A
31	27	77.1	160	4	US-09-270-767-50924	Sequence 50924, A
32	27	77.1	199	4	US-09-710-279-1170	Sequence 1170, Ap
33	27	77.1	218	4	US-09-252-591A-21404	Sequence 21404, Ap
34	27	77.1	225	2	US-08-738-462-2	Sequence 2, Appl
35	27	77.1	225	5	PCT-US94-07587-2	Sequence 2, Appl
36	27	77.1	238	4	US-08-716-301-10	Sequence 10, Appl
37	27	77.1	238	4	US-09-874-585D-44	Sequence 44, Appl
38	27	77.1	255	4	US-09-248-796A-18600	Sequence 18600, A
39	27	77.1	270	3	US-09-134-001C-5561	Sequence 5561, Ap
40	27	77.1	341	3	US-09-134-001C-4077	Sequence 4077, Ap
41	27	77.1	350	4	US-09-949-016-6257	Sequence 6257, Ap
42	27	77.1	366	4	US-09-248-796A-17059	Sequence 17059, A
43	27	77.1	370	4	US-09-949-016-7390	Sequence 7390, Ap
44	27	77.1	374	3	US-09-134-001C-2939	Sequence 2939, Ap
45	27	77.1	379	4	US-09-107-532A-5519	Sequence 5519, Ap

ALIGNMENTS

```
RESULT 1
US-09-198-452A-739
Sequence 739, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment,
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev.
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 739
LENGTH: 282
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-739

Query Match      88.6%; Score 31; DB 4; Length 282;
Best Local Similarity 71.4%; Pred. No. 39;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 EINHETD 7
Db      75 EVNVED 81

RESULT 2
US-09-438-185A-699
Sequence 699, Application US/09438185A
Patent No. 6822071
GENERAL INFORMATION:
APPLICANT: Stephens, Richard
APPLICANT: Mitchell, Wayne
APPLICANT: Kalman, Sue
APPLICANT: Davis, Ronald
TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
FILE REFERENCE: 018941-000411US
CURRENT APPLICATION NUMBER: US/09/438,185A
CURRENT FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/108,279
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 60/128,606
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 1074
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 699
LENGTH: 287
```

TYPE: PRT
 ORGANISM: Chlamydia pneumoniae
 FEATURE:
 OTHER INFORMATION: CPN0697
 US-09-438-185A-699

Query Match
 Best Local Similarity 88.6%; Score 31; DB 4; Length 287;
 Matches 5; Conservativity 71.4%; Pred. No. 40;
 Mismatches 2; Indels 0; Gaps 0;
 QY 1 EINLETD 7
 DB 80 EVNLETD 86

RESULT 3
 US-09-270-767-36607
 Sequence 36607, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 36607
 LENGTH: 146
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-270-767-36607

Query Match
 Best Local Similarity 85.7%; Score 30; DB 4; Length 146;
 Matches 6; Conservativity 100.0%; Pred. No. 30;
 Mismatches 0; Indels 0; Gaps 0;
 QY 2 EINLETD 7
 DB 34 INLETD 39

RESULT 4
 US-09-270-767-51824
 Sequence 51824, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 51824
 LENGTH: 146
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-270-767-51824

Query Match
 Best Local Similarity 85.7%; Score 30; DB 4; Length 146;
 Matches 6; Conservativity 100.0%; Pred. No. 30;
 Mismatches 0; Indels 0; Gaps 0;
 QY 2 EINLETD 7
 DB 34 INLETD 39

RESULT 5
 US-08-991-761A-8
 Sequence 8, Application US/08991761A
 Patent No. 6576609

GENERAL INFORMATION:
 APPLICANT: Soff, Gerald
 APPLICANT: Gately, Stephen
 APPLICANT: Twardowski, Przemyslaw
 TITLE OF INVENTION: "Methods and Compositions for Generating
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross P.C.
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: CO
 COUNTRY: USA
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/991,761A
 FILING DATE:
 CLASSIFICATION: 1642
 ATTORNEY/AGENT INFORMATION:
 NAME: Crook, Wamell M.
 REGISTRATION NUMBER: 31,071
 REFERENCE/DOCKET NUMBER: 3501-16-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 333 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-991-761A-8

Query Match
 Best Local Similarity 85.7%; Score 30; DB 4; Length 333;
 Matches 5; Conservativity 71.4%; Pred. No. 79;
 Mismatches 2; Indels 0; Gaps 0;
 QY 1 EINLETD 7
 DB 165 EVNLETD 171

RESULT 6
 US-09-543-681A-5035
 Sequence 5035, Application US/09543681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 FILE REFERENCE: 2709,1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 5035
 LENGTH: 964
 TYPE: PRT
 ORGANISM: Proteus mirabilis
 US-09-543-681A-5035

Query Match
 Best Local Similarity 85.7%; Score 30; DB 4; Length 964;
 Matches 6; Conservativity 85.7%; Pred. No. 2.7e+02;
 Mismatches 1; Indels 0; Gaps 0;
 QY 1 EINLETD 7

Db 415 EINLETD 421

RESULT 7
US-09-248-796A-16599
; Sequence 16599, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248, 796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16599
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16599

Query Match 82.9%; Score 29; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 6
Db 103 EINLETD 108

RESULT 8
US-09-134-001C-4971
; Sequence 4971, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4971
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4971

Query Match 82.9%; Score 29; DB 3; Length 208;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EINLETD 7
Db 74 EINLETD 80

RESULT 9
US-09-902-540-15089
; Sequence 15089, Application US/09902540
; Patent No. 6831447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.
; APPLICANT: Miegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15089
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-15089

Query Match 82.9%; Score 29; DB 4; Length 210;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EINLETD 7
Db 69 EVNCETD 75

RESULT 10
US-09-248-796A-22971
; Sequence 22971, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248, 796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22971
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22971

Query Match 82.9%; Score 29; DB 4; Length 247;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLETD 6
Db 129 EINLETD 134

RESULT 11
US-09-710-279-752
; Sequence 752, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 752
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-09-710-279-752

Query Match
Best Local Similarity 82.9%; Score 29; DB 4; Length 273;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLET 6
DB 247 EINLET 252

RESULT 12
US-09-710-279-2668
Sequence 2668, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2668
LENGTH: 307
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-2668

Query Match
Best Local Similarity 82.9%; Score 29; DB 4; Length 307;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLET 6
DB 281 EINLET 286

RESULT 13
US-09-248-796A-18201
Sequence 18201, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Kelch Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 18201
LENGTH: 314
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: UNSURE
LOCATION: (281), (284), (287), (294), (300), (301), (309), (311), (312), (313), (314)
OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-18201

Query Match 82.9%; Score 29; DB 4; Length 314;

Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 INLETD 7
DB 163 UNLETD 168

RESULT 14
US-09-134-001C-5553
Sequence 5553, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5553
LENGTH: 319
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5553

Query Match
Best Local Similarity 82.9%; Score 29; DB 3; Length 319;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EINLET 6
DB 293 EINLET 298

RESULT 15
US-09-583-110-4370
Sequence 4370, Application US/09583110
Patent No. 6699703
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
FILE REFERENCE: PAT000-07A
CURRENT APPLICATION NUMBER: US/09/583,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4370
LENGTH: 346
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4370

Query Match
Best Local Similarity 82.9%; Score 29; DB 4; Length 346;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EINLETD 7
DB 75 EVNALETD 81

Search completed: August 26, 2005, 01:09:50

Job time : 31 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using bw model

Run on: August 26, 2005, 01:01:32 ; Search time 83 Seconds
(without alignments)
33.138 Million cell updates/sec

Title: US-10-726-967A-57
Perfect score: 35
Sequence: 1 EINLETD 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	7	US-10-726-967A-57	Sequence 57, Appl
2	35	100.0	433	US-10-726-967A-84	Sequence 84, Appl
3	33	94.3	17	US-10-726-967A-56	Sequence 56, Appl
4	33	94.3	433	US-10-726-967A-81	Sequence 81, Appl
5	31	88.6	7	US-10-726-967A-58	Sequence 58, Appl
6	31	88.6	282	US-10-289-762-739	Sequence 739, Appl
7	31	88.6	282	US-10-282-122A-54697	Sequence 54697, A
8	31	88.6	282	US-10-282-122A-55050	Sequence 55050, A
9	30	85.7	252	US-10-988-943-4	Sequence 4, Appl
10	30	85.7	403	US-10-369-493-1103	Sequence 1103, Ap
11	30	85.7	506	US-10-225-068-180	Sequence 180, App

12	30	85.7	506	15	US-10-374-780A-366	Sequence 366, App
13	30	85.7	506	17	US-10-225-068-180	Sequence 180, App
14	30	85.7	767	15	US-10-094-749-2369	Sequence 2369, Ap
15	30	85.7	929	15	US-10-369-493-11264	Sequence 11264, A
16	30	85.7	958	15	US-10-282-122A-69140	Sequence 69140, A
17	30	85.7	1400	16	US-10-473-576-21	Sequence 21, Appl
18	29	82.9	64	16	US-10-425-115-358656	Sequence 358656,
19	29	82.9	85	9	US-09-815-242-12315	Sequence 12315, A
20	29	82.9	159	15	US-10-424-599-24062	Sequence 24062,
21	29	82.9	176	15	US-10-335-977-6151	Sequence 6151, Ap
22	29	82.9	184	15	US-10-335-977-6152	Sequence 6152, Ap
23	29	82.9	195	16	US-10-767-701-34598	Sequence 34598, A
24	29	82.9	198	13	US-10-067-989-7	Sequence 7, Appl
25	29	82.9	208	18	US-10-724-972A-4402	Sequence 4402, Ap
26	29	82.9	210	17	US-10-732-923-9054	Sequence 9054, Ap
27	29	82.9	211	15	US-10-424-599-24062	Sequence 193342,
28	29	82.9	221	17	US-10-732-923-9055	Sequence 9055, Ap
29	29	82.9	224	17	US-10-732-923-9076	Sequence 9076, Ap
30	29	82.9	272	17	US-10-741-600-1607	Sequence 1607, Ap
31	29	82.9	275	15	US-10-282-122A-53973	Sequence 53973, A
32	29	82.9	280	15	US-10-335-977-7108	Sequence 7108, Ap
33	29	82.9	284	15	US-10-282-122A-65204	Sequence 65204, A
34	29	82.9	284	15	US-10-282-122A-65681	Sequence 65681, A
35	29	82.9	287	15	US-10-425-114-36754	Sequence 36754, A
36	29	82.9	287	16	US-10-425-115-247023	Sequence 247023,
37	29	82.9	290	9	US-09-815-242-5743	Sequence 5743, Ap
38	29	82.9	292	15	US-10-282-122A-51031	Sequence 51031, A
39	29	82.9	292	15	US-10-282-122A-61133	Sequence 61133, A
40	29	82.9	292	15	US-10-282-122A-70936	Sequence 70936, A
41	29	82.9	292	15	US-10-282-122A-71459	Sequence 71459, A
42	29	82.9	293	9	US-09-815-242-12742	Sequence 12742, A
43	29	82.9	293	15	US-10-282-122A-43901	Sequence 43901, A
44	29	82.9	293	17	US-10-857-625-636	Sequence 636, App
45	29	82.9	294	15	US-10-282-122A-60812	Sequence 60812, A

ALIGNMENTS

RESULT 1
US-10-726-967A-57
; Sequence 57, Application US/10726967A
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OR INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Engineered BACE1 autoproteolysis site
US-10-726-967A-57

Query Match 100.0%; Score 35; DB 17; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 1 EINLETD 7
|||||
Db 1 EINLETD 7

RESULT 2
US-10-726-967A-84
; Sequence 84, Application US/10726967A
; Publication No. US20050074456A1

```

; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Soluble human probACE1 containing an engineered E1NL site
; OTHER INFORMATION: starting at position 21
US-10-726-967a-84

```

```

Query Match          100.0%; Score 35; DB 17; Length 433;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 E1NLETD 7
    |||||
Db 21 E1NLETD 27

```

```

RESULT 3
US-10-726-967a-56
; Sequence 56, Application US/10726967A
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Engineered BACE1 autoproteolysis site
US-10-726-967a-56

```

```

Query Match          94.3%; Score 33; DB 17; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.6e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 E1NLETD 7
    |||||
Db 1 E1NLETD 7

```

```

RESULT 4
US-10-726-967a-81
; Sequence 81, Application US/10726967A
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 433
; TYPE: PRT

```

```

; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Soluble human probACE1
US-10-726-967a-81

```

```

Query Match          94.3%; Score 33; DB 17; Length 433;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 E1NLETD 7
    |||||
Db 21 E1NLETD 27

```

```

RESULT 5
US-10-726-967a-58
; Sequence 58, Application US/10726967A
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Engineered BACE1 autoproteolysis site
US-10-726-967a-58

```

```

Query Match          88.6%; Score 31; DB 17; Length 7;
Best Local Similarity 85.7%; Pred. No. 1.6e+06;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 E1NLETD 7
    |||||
Db 1 E1NLETD 7

```

```

RESULT 6
US-10-289-762-739
; Sequence 739, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griffls, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments,
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 739
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-739

```

```

Query Match          88.6%; Score 31; DB 15; Length 282;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 E1NLETD 7
    |||||
Db 75 E1NLETD 81

```

```

RESULT 7

```

```

US-10-282-122A-54697
; Sequence 54697, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54697
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-282-122A-54697

Query Match      88.6%; Score 31; DB 15; Length 282;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 INLETD 7
|:|:|
Db      75 EVNVEITD 81

RESULT 8
US-10-282-122A-55050
; Sequence 55050, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert

```

```

; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55050
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-282-122A-55050

Query Match      88.6%; Score 31; DB 15; Length 282;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 INLETD 7
|:|:|
Db      75 EVNVEITD 81

RESULT 9
US-10-988-943-4
; Sequence 4, Application US/10988943
; Publication No. US20050176085A1
; GENERAL INFORMATION:
; APPLICANT: Center for Genetic Engineering and Biotechnology
; TITLE OF INVENTION: METHOD OF SELECTIVE PEPTIDE ISOLATION FOR THE
; TITLE OF INVENTION: IDENTIFICATION AND QUANTITATIVE ANALYSIS OF PROTEINS IN
; FILE REFERENCE: Proteomics CU2003-269
; CURRENT APPLICATION NUMBER: US/10/988,943
; PRIOR FILING DATE: 2004-11-15
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Neisseria meningitidis (group B)
US-10-988-943-4

Query Match      85.7%; Score 30; DB 18; Length 252;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 INLETD 7
|:|:|
Db      64 INLETD 69

```

```
RESULT 10
US-10-369-493-1103
; Sequence 1103, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1103
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Mechanobacterium thermotrophicum
US-10-369-493-1103
```

```
Query Match
Best Local Similarity 85.7%; Score 30; DB 15; Length 403;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EINLETD 7
Db 282 EINLEND 288
```

```
RESULT 11
US-10-225-068-180
; Sequence 180, Application US/10225068
; Publication No. US20030217383A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omaria
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Broun, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 514442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (214)...(287)
; OTHER INFORMATION: Conserved domain
US-10-225-068-180
```

```
Query Match
Best Local Similarity 85.7%; Score 30; DB 15; Length 506;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EINLETD 7
Db 84 EISLETD 90
```

```
RESULT 12
US-10-374-780A-366
; Sequence 366, Application US/10374780A
; Publication No. US20040019927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Dubell III, Arnold T.
; APPLICANT: Pineda, Omaria
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 366
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G1919 (conserved domain in AA coordinates:214-287)
US-10-374-780A-366
```

```
Query Match
Best Local Similarity 85.7%; Score 30; DB 15; Length 506;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 EINLETD 7
Db 84 EISLETD 90
```

```
RESULT 13
US-10-225-068-180
```

```
; Sequence 180, Application US/10225068
; Publication No. US20050120408A9
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Adam, Luc J.
; APPLICANT: Dubell, Arnold T.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pineda, Omalta
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Biron, Pierre E.
; TITLE OF INVENTION: STRESS-RELATED POLYNUCLEOTIDES AND
; FILE REFERENCE: 51442002040
; CURRENT APPLICATION NUMBER: US/10/225,068
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (214)...(287)
; OTHER INFORMATION: Conserved domain
; US-10-225-068-180

Query Match      85.7%; Score 30; DB 17; Length 506;
Best Local Similarity 85.7%; Pred. No. 9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EINLETD 7
       ||:||||
Db      84 EISLETD 90
```

```
RESULT 14
; Sequence 2369, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
```

```
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2369
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-094-749-2369
```

```
Query Match      85.7%; Score 30; DB 15; Length 767;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 EINLETD 7
       ||:||||
Db      467 DINLETD 473
```

```
RESULT 15
; Sequence 11264, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11264
; LENGTH: 929
; TYPE: PRT
; ORGANISM: Methanosarcina mazei
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(929)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-369-493-11264
```

```
Query Match      85.7%; Score 30; DB 15; Length 929;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2 INLETD 7
       ||:||||
Db      594 INLETD 599
```

```
Search completed: August 26, 2005, 01:11:18
Job time : 84 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 26, 2005, 01:08:52 ; Search time 2605 Seconds
(without alignments)
130.206 Million cell updates/sec

Title: US-10-726-967a-57
Perfect score: 35
Sequence: 1 EINFRTD 7

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-Q=/cgn2_1/USFTO.spool_h/US10726967/runat_25082005_170153_283/app_query.fasta.1.199
-DB=GenEmbl -OPMT=faetap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10726967 @CGN 1.1 3731 @runat_25082005_170153_283 -NCPu=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSRBDLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	35	100.0	347	6	AR148046
C 2	35	100.0	347	6	AR219770
C 3	35	100.0	347	6	AR350868
C 4	35	100.0	347	6	AR433244

C 5	35	100.0	347	6	AX316888	AX316888 Sequence
C 6	35	100.0	347	6	BD084421	BD084421 Composite
C 7	35	100.0	384	6	AR508950	AR508950 Sequence
C 8	35	100.0	1729	3	AK173524	AK173524 Clona int
C 9	35	100.0	81953	2	AC016835	AC016835 Homo sapi
C 10	35	100.0	93919	2	AC002415	AC002415 Homo sapi
C 11	35	100.0	104059	5	BX901938	BX901938 Zebrafish
C 12	35	100.0	110000	2	AC128488_2	Continuation (3 of
C 13	35	100.0	11560	9	AC073905	AC073905 Homo sapi
C 14	35	100.0	114344	2	AL445227	AL445227 Homo sapi
C 15	35	100.0	119074	9	AL359687	AL359687 Human DNA
C 16	35	100.0	121419	5	BX324226	BX324226 Zebrafish
C 17	35	100.0	148726	5	BX005402	BX005402 Zebrafish
C 18	35	100.0	149749	9	AL445236	AL445236 Human DNA
C 19	35	100.0	159397	9	AC068544	AC068544 Homo sapi
C 20	35	100.0	160429	2	AC136950	AC136950 Homo sapi
C 21	35	100.0	164223	10	AC127228	AC127228 Mus muscu
C 22	35	100.0	172064	9	AC146460	AC146460 Pan trogl
C 23	35	100.0	182846	2	AC068223	AC068223 Homo sapi
C 24	35	100.0	192060	10	AC121610	AC121610 Mus muscu
C 25	35	100.0	203241	9	AC008282	AC008282 Homo sapi
C 26	35	100.0	213499	10	AC110541	AC110541 Mus muscu
C 27	35	100.0	216041	10	AC137585	AC137585 Mus muscu
C 28	35	100.0	217068	2	AC079434	AC079434 Mus muscu
C 29	35	100.0	244130	2	AC125658	AC125658 Rattus no
C 30	35	100.0	249936	2	AC123096	AC123096 Rattus no
C 31	35	100.0	250731	2	AC105473	AC105473 Rattus no
C 32	35	100.0	252314	2	AC114502	AC114502 Rattus no
C 33	35	100.0	301200	1	AP005374	AP005374 Thermosyn
C 34	35	100.0	335593	14	AF204951	AF204951 Ectocarpu
C 35	35	100.0	335593	14	AF204951	AF204951 Ectocarpu
C 36	35	97.1	653	8	AY144859	AY144859 Saccharom
C 37	34	97.1	23088	10	AF330047	AF330047 Mus muscu
C 38	34	97.1	49768	9	AL365180	AL365180 Human DNA
C 39	34	97.1	147211	9	HS37816	HS37816 Homo sapi
C 40	34	97.1	148852	9	AC005878	AC005878 Cltb.255
C 41	34	97.1	165116	2	AL512431	AL512431 Homo sapi
C 42	34	97.1	176493	10	AC102661	AC102661 Mus muscu
C 43	34	97.1	186315	2	AC110188	AC110188 Mus muscu
C 44	34	97.1	198984	2	AC093452	AC093452 Mus muscu
C 45	34	97.1	199932	2	AC114521	AC114521 Rattus no

ALIGNMENTS

RESULT 1
AR148046/c
LOCUS AR148046 347 bp DNA
DEFINITION Sequence 225 from patent US 6225054.
ACCESSION AR148046
VERSION AR148046.1 GI:15112136
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 347)
AUTHORS Frudakis,T.N., Smith,J.M. and Reed,S.G.
TITLE Compositions and methods for the treatment and diagnosis of breast
JOURNAL Patent: US 6225054-A 225 01-MAY-2001;
FEATURES
source location/Qualifiers
1..347
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 7.6
Score: 35.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
Length: 347
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-726-967A-57 (1-7) x AR148046 (1-347)

Qy 1 Glu1leasleuGluThrAsp 7
 Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 2
 LOCUS AR219770/c 347 bp DNA linear PAT 26-SEP-2002

DEFINITION Sequence 225 from patent US 6423496.
 ACCESSION AR219770
 VERSION AR219770.1 GI:23232948
 KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 347)

AUTHORS Frudakis,T.N., Smith,J.M. and Reed,S.G.
 TITLE Compositions and methods for the treatment and diagnosis of breast cancer
 JOURNAL Patent: US 6423496-A 225 23-JUL-2002;
 FEATURES Location/Qualifiers
 source 1..347
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.6 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-726-967A-57 (1-7) x AR219770 (1-347)

Qy 1 Glu1leasleuGluThrAsp 7
 Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 3
 LOCUS AR350868/c 347 bp DNA linear PAT 17-AUG-2003

DEFINITION Sequence 225 from patent US 6586570.
 ACCESSION AR350868
 VERSION AR350868.1 GI:33752508
 KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 347)
 AUTHORS Frudakis,T.N., Reed,S.G., Smith,J.M. and Misher,L.
 TITLE Compositions and methods for the treatment and diagnosis of breast cancer
 JOURNAL Patent: US 6586570-A 225 01-JUL-2003;
 FEATURES Location/Qualifiers
 source 1..347
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.6 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-726-967A-57 (1-7) x AR350868 (1-347)

Qy 1 Glu1leasleuGluThrAsp 7
 Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 4
 LOCUS AR433244/c 347 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 225 from patent US 6656480.
 ACCESSION AR433244
 VERSION AR433244.1 GI:40196026
 KEYWORDS

SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 347)

AUTHORS Retter,M.W. and Dillon,D.C.
 TITLE Compositions and methods for the treatment and diagnosis of breast cancer
 JOURNAL Patent: US 6656480-A 225 02-DEC-2003;
 FEATURES Location/Qualifiers
 source 1..347
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.6 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-726-967A-57 (1-7) x AR433244 (1-347)

Qy 1 Glu1leasleuGluThrAsp 7
 Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 5
 LOCUS AX316888/c 347 bp DNA linear PAT 14-DEC-2001

DEFINITION Sequence 225 from Patent WO0190152.
 ACCESSION AX316888
 VERSION AX316888.1 GI:17899969
 KEYWORDS

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Retter,M.W., Wang,A., Skeiky,Y.A., Harlocker,S.L. and Day,C.H.
 TITLE Compositions and methods for the therapy and diagnosis of breast cancer
 JOURNAL Patent: WO 0190152-A 225 29-NOV-2001;
 FEATURES Location/Qualifiers
 source 1..347
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 7.6 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-726-967A-57 (1-7) x AX316888 (1-347)

QY 1 Glut1easnlenglurhrasp 7
Db 114 GAAATTAATTGAAACGAT 94

RESULT 6
LOCUS BD084421/c 347 bp DNA linear PAT 27-AUG-2002
DEFINITION Compositions and methods for the treatment and diagnosis of breast cancer.
ACCESSION BD084421
VERSION BD084421.1 GI:22630031
KEYWORDS JP 2001521384-A/214.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 347)
AUTHORS Frudakis,T.N., Smith,J.M. and Reed,S.G.
TITLE Compositions and methods for the treatment and diagnosis of breast cancer.
JOURNAL Patent: JP 2001521384-A 214 06-NOV-2001;
CORIXA CORP

COMMENT OS Unidentified
PN JP 2001521384-A/214
PD 06-NOV-2001
PF 09-APR-1998 JP 1998543059
PR 09-APR-1997 US 08/838762,11-DEC-1997 US 08/991789 PI
TONY N FRUDAKIS,JOHN M SMITH,STEVEN G REED
PC C07K14/47,C07K14/82,C07K14/15,C12Q1/68,G01N3/574,A61K38/17,
A61K39/00
CC Strandedness: Single;
CC Topology: Linear;
CC Compositions and methods for the treatment and diagnosis of
breast cancer
Citation/Qualifiers
FT source 1.347
/location/Qualifiers
1.347
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Alignment Scores:
Pred. No.: 7 6 Length: 347
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-726-967a-57 (1-7) x BD084421 (1-347)

QY 1 Glut1easnlenglurhrasp 7
Db 114 GAAATTAATTGAAACGAT 94

RESULT 7
LOCUS AR508950/c 384 bp DNA linear PAT 22-SEP-2004
DEFINITION Sequence 13910 from patent US 6703491.
ACCESSION AR508950
VERSION AR508950.1 GI:52444425
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 384)
AUTHORS Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Liang,H.L.,
Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
TITLE Drosophila sequences
JOURNAL Patent: US 6703491-A 13910 09-MAR-2004;

FEATURES
source Location/Qualifiers
1.384
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 8 52 Length: 384
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-726-967a-57 (1-7) x AR508950 (1-384)

QY 1 Glut1easnlenglurhrasp 7
Db 93 GAAATCAATTGGAACCGAT 73

RESULT 8
LOCUS AK173524 1729 bp mRNA linear INV 01-SEP-2004
DEFINITION Ciona intestinalis cDNA, clone:ciad013b07, full insert sequence.
ACCESSION AK173524
VERSION AK173524.1 GI:51774018
KEYWORDS FLI cDNA.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Ciona.

REFERENCE
AUTHORS 1 Satou,Y., Yamada,L., Mochizuki,Y., Takatori,N., Kawashima,T.,
Sasaki,A., Hamaguchi,M., Awazu,S., Yagi,K., Sasaki,Y.,
Makayama,A., Ishikawa,H., Inaba,K. and Satoh,N.
A cDNA resource from the basal chordate Ciona intestinalis
Genesis 33 (4), 153-154 (2002)
MEDLINE 22191024
PUBMED 12203911

REFERENCE
AUTHORS 2 Satou,Y. and Satoh,N.
A cDNA resource from the basal chordate Ciona intestinalis
Published Only in Database (2004)

JOURNAL
REFERENCE 3 (bases 1 to 1729)
AUTHORS Satou,Y.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2004) Yutaka Satou, Kyoto University, Department
of Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:yutaka@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,
Fax:81-75-705-1113)
Ciona intestinalis cDNA Project (URL:
http://ghost.zool.kyoto-u.ac.jp/index1.html).

COMMENT
http://ghost.zool.kyoto-u.ac.jp/index1.html).

FEATURES
source Location/Qualifiers
1.1729
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad013b07"

ORIGIN
Alignment Scores:
Pred. No.: 47 3 Length: 1729
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-726-967a-57 (1-7) x AK173524 (1-1729)

QY 1 Glut1easnlenglurhrasp 7
Db 324 GAAATTAACCTTGAAACTGAT 344

RESULT 9
AC016835/c
LOCUS AC016835 81953 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens clone RP11-1P5, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC016835
VERSION AC016835.2 GI:9119883
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE 1 (bases 1 to 81953)
JOURNAL Homo sapiens, clone RP11-1P5
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 81953)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Boukhalter, B., Brown, A., Caciule, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Gallegan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lien, C., Locke, K., Macdonald, P., Marquis, N., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teefaye, S., Tirelli, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 11, 2000 this sequence version replaced gi:5539374.
All repeats were identified using RepeatMasker:
Smt, A. F. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: W1BR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L2520

Center clone name: 1_P_5

* NOTE: This record contains 91 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 769: contig of 769 bp in length
* 770 869: gap of 100 bp
* 870 1691: contig of 822 bp in length
* 1692 1791: gap of 100 bp
* 1792 2590: contig of 799 bp in length
* 2591 2690: gap of 100 bp
* 2691 3502: contig of 812 bp in length
* 3503 3602: gap of 100 bp
* 3603 4403: contig of 801 bp in length
* 4404 4503: gap of 100 bp
* 4504 5271: contig of 768 bp in length
* 5272 5371: gap of 100 bp
* 5372 6160: contig of 789 bp in length
* 6161 6260: gap of 100 bp
* 6261 7048: contig of 788 bp in length

7049 7148: gap of 100 bp
* 7149 7149: contig of 797 bp in length
* 7946 8045: gap of 100 bp
* 8046 8864: contig of 815 bp in length
* 8865 8965: gap of 100 bp
* 8966 9768: contig of 804 bp in length
* 9769 9868: gap of 100 bp
* 9869 10680: contig of 811 bp in length
* 10680 10779: gap of 100 bp
* 10780 11545: contig of 766 bp in length
* 11546 11645: gap of 100 bp
* 11646 12460: contig of 815 bp in length
* 12461 12560: gap of 100 bp
* 12561 13372: contig of 812 bp in length
* 13373 13472: gap of 100 bp
* 13473 14275: contig of 803 bp in length
* 14276 14376: gap of 100 bp
* 14377 15173: contig of 798 bp in length
* 15174 15273: gap of 100 bp
* 15274 16058: contig of 785 bp in length
* 16059 16158: gap of 100 bp
* 16159 16953: contig of 795 bp in length
* 16954 17053: gap of 100 bp
* 17054 17840: contig of 787 bp in length
* 17841 17940: gap of 100 bp
* 17941 18750: contig of 810 bp in length
* 18751 18850: gap of 100 bp
* 18851 19656: contig of 806 bp in length
* 19657 19756: gap of 100 bp
* 19757 20561: contig of 805 bp in length
* 20562 21456: gap of 100 bp
* 21457 21557: contig of 795 bp in length
* 21558 22382: contig of 826 bp in length
* 22383 22482: gap of 100 bp
* 22483 23306: contig of 824 bp in length
* 23307 23406: gap of 100 bp
* 23407 24217: contig of 811 bp in length
* 24218 24317: gap of 100 bp
* 24319 25115: contig of 799 bp in length
* 25117 25216: gap of 100 bp
* 25217 25999: contig of 783 bp in length
* 26000 26099: gap of 100 bp
* 26100 26879: contig of 780 bp in length
* 26880 26979: gap of 100 bp
* 26980 27773: contig of 794 bp in length
* 27774 27873: gap of 100 bp
* 27874 28668: contig of 795 bp in length
* 28669 28768: gap of 100 bp
* 28769 29592: contig of 824 bp in length
* 29593 29692: gap of 100 bp
* 29693 30523: contig of 831 bp in length
* 30524 30623: gap of 100 bp
* 30624 31423: contig of 800 bp in length
* 31424 31523: gap of 100 bp
* 31524 32321: contig of 798 bp in length
* 32322 32421: gap of 100 bp
* 32422 33244: contig of 823 bp in length
* 33245 33344: gap of 100 bp
* 33345 34172: contig of 828 bp in length
* 34173 34272: gap of 100 bp
* 34273 35079: contig of 807 bp in length
* 35080 35179: gap of 100 bp
* 35180 35967: contig of 788 bp in length
* 35968 36067: gap of 100 bp
* 36068 36846: contig of 779 bp in length
* 36847 37740: gap of 100 bp
* 37741 37840: gap of 100 bp
* 37841 38651: contig of 811 bp in length
* 38652 38751: gap of 100 bp
* 38752 39566: contig of 815 bp in length
* 39567 39666: gap of 100 bp

```

* 39667 40472: contig of 806 bp in length
* 40473 40572: gap of 100 bp
* 40573 41381: contig of 809 bp in length
* 41382 41481: gap of 100 bp
* 41482 42296: contig of 815 bp in length
* 42297 42396: gap of 100 bp
* 42397 43213: contig of 817 bp in length
* 43214 43313: gap of 100 bp
* 43314 44124: contig of 811 bp in length
* 44125 44224: gap of 100 bp
* 44225 45028: contig of 804 bp in length
* 45029 45128: gap of 100 bp
* 45129 45937: contig of 809 bp in length
* 45938 46037: gap of 100 bp
* 46038 46814: contig of 777 bp in length
* 46815 46914: gap of 100 bp
* 46915 47710: contig of 796 bp in length
* 47711 47810: gap of 100 bp
* 47811 48580: contig of 770 bp in length
* 48581 49465: contig of 785 bp in length
* 49466 49565: gap of 100 bp
* 49566 50377: contig of 812 bp in length
* 50378 50477: gap of 100 bp
* 50478 51260: contig of 783 bp in length
* 51261 51361: gap of 100 bp
* 51361 52243: contig of 783 bp in length
* 52244 53049: gap of 100 bp
* 53050 53149: gap of 100 bp
* 53150 53959: contig of 810 bp in length
* 53960 54059: gap of 100 bp
* 54060 54832: contig of 773 bp in length
* 54833 54932: gap of 100 bp
* 54933 55712: contig of 780 bp in length
* 55713 55812: gap of 100 bp
* 55813 56604: contig of 792 bp in length
* 56605 56704: gap of 100 bp
* 56705 57487: contig of 783 bp in length
* 57488 57587: gap of 100 bp
* 57588 58395: contig of 808 bp in length
* 58396 58495: gap of 100 bp
* 58496 59302: contig of 807 bp in length
* 59303 59402: gap of 100 bp
* 59403 60205: contig of 803 bp in length
* 60206 60305: gap of 100 bp
* 60306 61119: contig of 814 bp in length
* 61120 61219: gap of 100 bp
* 61220 62026: contig of 807 bp in length
* 62027 62126: gap of 100 bp
* 62127 62941: contig of 815 bp in length
* 62942 63041: gap of 100 bp
* 63042 63857: contig of 816 bp in length
* 63858 63957: gap of 100 bp
* 63958 64772: contig of 815 bp in length
* 64773 64872: gap of 100 bp
* 64873 65675: contig of 803 bp in length

```

Alignment Scores:

Pred. No.:	3.84e+03	Length:	81953
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-726-967a-57 (1-7) x AC016835 (1-81953)

Oy 1 GluileleleugluThraap 7
 Db 17159 GAATCAATCTTGAAACGAT 17139

RESULT 10

AC002415
 LOCUS 93919 bp DNA linear HTG 13-JUN-2002
 DEFINITION Homo sapiens chromosome X clone bwxdl42, *** SEQUENCING IN PROGRESS
 *** 6 unordered pieces.
 AC002415
 VERSION AC002415.2 GI:21405635
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 93919)
 Chen,E., Brownstein,B.H., States,D.J., Schlessinger,D. and
 Mazzarella,R.
 Direct Submission
 Unpublished (1997)
 2 (bases 1 to 93919)
 Brownstein,B.H., States,D.J. and Mazzarella,R.
 Direct Submission
 Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232,
 Washington University School of Medicine, 4566 Scott Avenue, St.
 Louis, MO 63110, USA
 On Jun 13, 2002 this sequence version replaced gi:2323256.
 Current status of this project is available at:
 'http://genome.wustl.edu/cgm/seq_projects.html'
 Submitted by:
 Ellison Chen,
 Advanced Center for Genetic Technology,
 Applied Biosystems Division of Perlin Elmer Corp.,
 850 Lincoln Center Drive,
 Foster City, CA 94404 USA
 e-mail: ellison@genseq.apltdio.com

and

Buddy Brownstein,
 Center for Genetics in Medicine,
 Washington University School of Medicine, Box 8232
 4566 Scott Avenue,
 St. Louis, MO 63110, USA
 e-mail: buddy@genetics.wustl.edu

and

David J. States,
 Institute for Biomedical Computing
 Washington University in St. Louis
 700 South Euclid Ave.
 St. Louis, MO 63108 USA
 e-mail: states@ibc.wustl.edu.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

* 1 9560: contig of 9560 bp in length
* 9561 9660: gap of 100 bp
* 9661 25227: contig of 15567 bp in length
* 25228 25327: gap of 100 bp
* 25328 45875: contig of 20548 bp in length
* 45876 45975: gap of 100 bp
* 45976 73699: contig of 27724 bp in length
* 73700 73799: gap of 100 bp
* 73800 82475: contig of 8676 bp in length
* 82476 82575: gap of 100 bp
* 82576 93919: contig of 11344 bp in length.

```

FEATURES
 source 1..93919
 /organism="Homo sapiens"
 /mol_type="genomic DNA"

/db_xref="taxon:9606"
/chromosome="X"
/clone="bmx142"

ORIGIN

Alignment Scores:
Pred. No.: 4.48e+03 Length: 93919
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AC002415 (1-93919)

OY 1 Glut1aenLeuGluThrAsp 7
Db 68521 GAATCATCTTGAAACAGAT 68541

RESULT 11

LOCUS BX901938 104059 bp DNA linear VRT 01-SEP-2004
DEFINITION zebrafish DNA sequence from clone CH211-254D18 in linkage group 1,
complete sequence.

ACCESSION BX901938
VERSION BX901938.5 GI:51870391
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 104059)
AUTHORS Johnson, C.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 3, 2004 this sequence version replaced gi:51592226.

COMMENT -----
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information
on the WormPep database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormPep
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhifeng Bao and Sean Bddy, submitted), and those
beginning 'dr' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml

CH211-254D18 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.
Location/Qualifiers
1. 104059

FEATURES
source /organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-254D18"
/clone_11b="CHORI-211"

ORIGIN

Alignment Scores:
Pred. No.: 5.04e+03 Length: 104059
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-726-967a-57 (1-7) x BX901938 (1-104059)

OY 1 Glut1aenLeuGluThrAsp 7
Db 68540 GAGATTAACCTTGAAACAGAT 68560

RESULT 12

LOCUS AC128488_2/c
WPCOMMENT Sequence split into 4 fragments
Fragment Name Begin End
AC128488_1 1 110000
AC128488_2 100001 210000
AC128488_3 200001 310000
AC128488_4 402823

Continuation (3 of 4) of AC128488 from base 200001 (AC128488 Rattus norvegicus clone CH2
Alignment Scores:
Pred. No.: 5.37e+03 Length: 110000
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AC128488_2 (1-110000)

OY 1 Glut1aenLeuGluThrAsp 7
Db 21109 GAATTAACCTTGAAACAGAT 21089

RESULT 13
LOCUS AC073905/c 111560 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-745H15 from 2, complete sequence.
ACCESSION AC073905
VERSION AC073905.3 GI:15431265
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 111560)

AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 98063792
PubMed 9847074

REFERENCE 2 (bases 1 to 111560)
Tomlinson, C., Kozlowski, A. and Phillips, A.
TITLE The sequence of Homo sapiens BAC clone RP11-745H15
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 111560)
Waterston, R.H.

TITLE Direct Submission
JOURNAL Submitted (02-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 111560)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (05-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 111560)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Sep 5, 2001 this sequence version replaced gi:15029499.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0745H15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frenken,B., Tatemio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-132H23, 2000 bp overlap; the clone sequenced to the right is RP11-108G19. Actual start of this clone is at base position 86031 of RP11-132H23; actual end is at base position 111560 of RP11-745H15.
Location/Qualifiers
1..111560
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-745H15"
/clone_1id="RPCI-11"
85..114
/rpt_family="(T)n"
385..730
/note="match to EST BE062373 (NID:g8407023) "
repeat_region
835..857

/rpt_family="AT_rich"
2807..2923
/rpt_family="GA-rich"
repeat_region
2925..3330
/rpt_family="MaLR"
repeat_region
3351..3637
/rpt_family="MIR"
repeat_region
3638..3934
/rpt_family="Alu"
repeat_region
3935..4050
/rpt_family="MIR"
repeat_region
4305..4599
/rpt_family="Alu"
repeat_region
4424..4457
/rpt_family="(A)n"
misc_feature
4966..5303
/note="match to EST BF931133 (NID:g12348457) "
repeat_region
5246..5562
/rpt_family="Alu"
repeat_region
5528..5575
/rpt_family="(CAA)n"
misc_feature
5648..5911
/note="match to EST BF831731 (NID:g12179795) "
repeat_region
5694..5786
/rpt_family="L2"
repeat_region
6900..7008
/rpt_family="MER1_type"
repeat_region
7297..7355
/rpt_family="L2"
repeat_region
7851..8043
/rpt_family="MaLR"
repeat_region
9463..10352
/rpt_family="ACHobo"
repeat_region
10571..10590
/rpt_family="(A)n"
repeat_region
10717..10963
/rpt_family="L2"
repeat_region
11064..11339
/rpt_family="L1"
repeat_region
11468..11719
/rpt_family="MIR"
repeat_region
12810..13134
/rpt_family="MaLR"
repeat_region
13355..13398
/rpt_family="Alu"
repeat_region
14367..14641
/rpt_family="Alu"
repeat_region
14623..14644
/rpt_family="AT_rich"
repeat_region
14642..14807
/rpt_family="Alu"
repeat_region
16106..16131
/rpt_family="(CA)n"
repeat_region
16285..16389
/rpt_family="MIR"
misc_feature
16564..16889
/note="match to EST AA326063 (NID:g1978318) "
repeat_region
17973..18605
/rpt_family="L1"
repeat_region
18696..18833
/rpt_family="Alu"
repeat_region
18846..18883
/rpt_family="(CA)n"
repeat_region
18906..19082
/rpt_family="(TA)n"
repeat_region
20563..20842
/rpt_family="Alu"
repeat_region
21386..22136
/rpt_family="L2"
repeat_region
22546..22837
/rpt_family="Alu"
repeat_region
22935..23119
/rpt_family="L2"

```
repeat_region 23127..23284 /rpt_family="MIR"
repeat_region 23414..23501 /rpt_family="r2"
repeat_region 23680..23719 /rpt_family="TG)n"
repeat_region 23878..23931 /rpt_family="(TA)n"
repeat_region 24361..24561 /rpt_family="L2"
repeat_region 24992..25605 /rpt_family="BRV1"
repeat_region 25677..25918 /rpt_family="L1"
repeat_region 25920..26701 /rpt_family="BRV1"
repeat_region 26912..26947 /rpt_family="AT_rich"
repeat_region 27836..27877 /rpt_family="(CAT)n"
repeat_region 28767..29026 /rpt_family="Alu"
repeat_region 30320..30895 /rpt_family="Achoho"
repeat_region 31931..31790 /rpt_family="Alu"
repeat_region 32268..32417 /rpt_family="MIR"
repeat_region 32733..33045 /rpt_family="Alu"
repeat_region 33268..33507 /rpt_family="L2"
repeat_region 33810..33966
```

Alignment Scores:

Pred. No.:	5.45e+03	Length:	111560
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-726-967a-57 (1-7) x AC073905 (1-111560)

Qy 1 GluIIeAmluengIuThraP 7
DB 95290 GAATTAATCTAGAACAGAT 95270

RESULT 14
AL445227 114344 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome X clone RP13-97115, 6 unordered pieces.
DEFINITION AL445227
ACCESSION AL445227 GI:11611395
VERSION AL445227.7 GI:11611395
KEYWORDS HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE
AUTHORS McLay, K.
TITLE Direct Submision
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Dec 9, 2000 this sequence version replaced gi:11493333.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
----- Project Information
Center project name: hb97115

----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 112482 bases at least Q40
Consensus quality: 113245 bases at least Q30
Consensus quality: 113597 bases at least Q20
Insert size: 113644; sum-of-contigs
Insert size: 122396; 9.8% error; agarose-fp
Quality coverage: 5.91x in Q20 bases; sum-of-contigs quality
coverage: 5.72x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
source
1..114344
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP13-97115"
/clone_id="RPC1-13.1"
1..13809
/note="assembly_fragment:00241
fragment_chain:1
clone_end:SP6
vector_side:left"
13910..51507
/note="assembly_fragment:00680
fragment_chain:1"
51608..65679
/note="assembly_fragment:00207
fragment_chain:1"
65780..69064
/note="assembly_fragment:00300"
69165..101518
/note="assembly_fragment:01024"
101619..114344
/note="assembly_fragment:01825
clone_end:T7
vector_side:right"

ORIGIN

Alignment Scores:

Pred. No.:	5.61e+03	Length:	114344
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-726-967a-57 (1-7) x AL445227 (1-114344)

Qy 1 GluIIeAmluengIuThraP 7
DB 48320 GAATTAATCTAGAACAGAT 48340

RESULT 15
AL159987 119024 bp DNA linear PRI 05-APR-2001
LOCUS Human DNA sequence from clone RP11-167P23 on chromosome X, complete
sequence.
ACCESSION AL159987 GI:13559998
VERSION AL159987.19
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 119024)
AUTHORS Howden, P.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, UK
COMMENT On Apr 6, 2001 this sequence version replaced gi:13374303.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SwissProt; Tr, TrEMBL; Wp, WormPeP; Information on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/chrX
RP11-167P23 is from the library RPCT-11.1 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-167P23. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-465B19 is at 118925 in this sequence. The true right end of clone RP1-47K8 is at 100 in this sequence.

FEATURES
source location/Qualifiers
1..119024
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-167P23"
/clone_1lb="RPCT-11.1"
1..1665
repeat_region
/note="HERV1 repeat: matches 1583..3249 of consensus"
1704..1743
/note="20 copies 2 mer ac 97% conserved"
1744..2134
repeat_region
/note="MUT2B repeat: matches 1..392 of consensus"
2381..2705
repeat_region
/note="AluJo repeat: matches 1..312 of consensus"
3061..3437
repeat_region
/note="THE1C repeat: matches 1..371 of consensus"
3445..3470
repeat_region
/note="13 copies 2 mer tg 92% conserved"
4743..4916
repeat_region

repeat_region
/note="MIR repeat: matches 7..195 of consensus"
5565..5682
misc_feature
/note="MIR repeat: matches 101..213 of consensus"
9092..9323
misc_feature
/note="match: GSS: Em:AQ980738"
9325..9686
misc_feature
/note="match: GSS: Em:AQ980738"
complement(9528..9691)
/note="match: STS: Em:G06170"
9709..14228
repeat_region
/note="L1PA6 repeat: matches 1483..6143 of consensus"
14447..14834
misc_feature
/note="match: GSS: Em:AQ414791"
15106..15549
misc_feature
/note="match: GSS: Em:AQ559891"
15638..15846
misc_feature
/note="match: STS: Em:G03827"
16677..16718
repeat_region
/note="MIR repeat: matches 110..151 of consensus"
17584..17872
repeat_region
/note="AluSx repeat: matches 2..311 of consensus"
complement(18344..18789)
/note="match: GSS: Em:AF157981"
complement(18346..18734)
/note="match: GSS: Em:AQ720285"
complement(18439..18734)
/note="match: GSS: Em:AQ374048"
18737..19166
misc_feature
/note="match: GSS: Em:AQ359652"
19368..19398
repeat_region
/note="MIR52C repeat: matches 1..31 of consensus"
19399..20564
repeat_region
/note="MIR52A repeat: matches 42..1194 of consensus"
20559..20626
repeat_region
/note="MIR52A repeat: matches 1684..1751 of consensus"
20828..21058
repeat_region
/note="AluJo repeat: matches 45..277 of consensus"
21379..21815
repeat_region
/note="MIR66B repeat: matches 108..484 of consensus"
23053..23121
repeat_region
/note="L1MA5 repeat: matches 6225..6233 of consensus"
23381..23495
repeat_region
/note="MIR41A repeat: matches 97..290 of consensus"
complement(23805..24380)
/note="match: GSS: Em:AQ507754"
complement(23964..24410)
/note="match: GSS: Em:AQ408337"
complement(23994..24399)
/note="match: GSS: Em:AQ508802"
complement(24134..24410)
/note="match: GSS: Em:AQ336100"
24485..25168
repeat_region
/note="L1PA3 repeat: matches 5475..6152 of consensus"
26042..26697
repeat_region
/note="L1M4 repeat: matches 2242..2926 of consensus"
26809..26927
repeat_region
/note="FLAM C repeat: matches 1..119 of consensus"
26940..27076
repeat_region
/note="L1M4 repeat: matches 4992..5137 of consensus"
27380..27584
repeat_region
/note="MIR repeat: matches 62..262 of consensus"
27946..28065
repeat_region
/note="MIR5A repeat: matches 68..188 of consensus"
28840..28923
repeat_region
/note="AluJ/FRAM repeat: matches 216..301 of consensus"
28943..29055
misc_feature
/note="FLAM C repeat: matches 3..115 of consensus"
complement(29034..29436)
/note="match: GSS: Em:BS2857"
complement(29901..30381)
/note="match: GSS: Em:AQ039266"
29989..30062
repeat_region
/note="HERV140 repeat: matches 5295..5367 of consensus"

Search completed: August 26, 2005, 02:01:51
Job time : 2655 secs

```
repeat_region      30081. .30112  
                    /note="16 copies 2 mer tg 100% conserved"  
repeat_region      30144. .30513  
                    /note="MLT2FB repeat: matches 1. .366 of consensus"  
misc_feature       31753. .32334  
                    /note="match: GSS: Em:AQ209396"  
repeat_region      31770. .31935  
                    /note="MIR repeat: matches 29. .200 of consensus"  
misc_feature       32266. .32756  
                    /note="match: GSS: Em:B47902"  
misc_feature       32364. .32727  
                    /note="match: GSS: Em:AQ107228"  
repeat_region      32412. .32777  
                    /note="L1P8 repeat: matches 5793. .6159 of consensus"  
repeat_region      32835. .33112  
                    /note="L2 repeat: matches 2255. .2522 of consensus"  
repeat_region      33291. .33584  
                    /note="MER39 repeat: matches 98. .395 of consensus"  
repeat_region      34330. .34453  
                    /note="FLAM C repeat: matches 1. .124 of consensus"  
repeat_region      34597. .34970  
                    /note="LTR3 repeat: matches 1. .413 of consensus"  
misc_feature       36254. .36581  
                    /note="match: GSS: Em:AQ389427"  
repeat_region      36267. .36573  
                    /note="A1usc repeat: matches 1. .309 of consensus"  
misc_feature       36730. .36730  
                    /note="match: GSS: Em:AQ181456"  
repeat_region      38336. .38377  
                    /note="MIR repeat: matches 110. .151 of consensus"  
misc_feature       39110. .39110  
                    /note="match: STS: Em:HS220XG1"  
repeat_region      41596. .41706  
                    /note="L1 repeat: matches 4663. .4779 of consensus"  
repeat_region      41765. .41821  
                    /note="L2 repeat: matches 1163. .1219 of consensus"  
repeat_region      41849. .42117  
                    /note="L2 repeat: matches 811. .1102 of consensus"  
repeat_region      42146. .42812  
                    /note="MLT1-INTERNAL repeat: matches 201. .885 of  
                    consensus"  
repeat_region      43013. .43410  
                    /note="MLT1-INTERNAL repeat: matches 920. .1322 of  
                    consensus"  
repeat_region      43499. .43559  
                    /note="TIGER1 repeat: matches 2288. .2359 of consensus"  
repeat_region      43560. .43799  
                    /note="L1PBI repeat: matches 5909. .6155 of consensus"  
repeat_region      43800. .44089  
                    /note="TIGER1 repeat: matches 2078. .2288 of consensus"  
repeat_region      44132. .44550  
                    /note="TIGER1 repeat: matches 1650. .2094 of consensus"  
repeat_region      44701. .45249  
                    /note="MER68A repeat: matches 1. .568 of consensus"  
repeat_region      45741. .46034  
                    /note="A1ujo repeat: matches 1. .293 of consensus"  
repeat_region      46311. .46342
```

Alignment Scores:

```
Pred. No.:      5.87e+03      Length:      119024  
Score:          35.00         Matches:      7  
Percent Similarity: 100.00%   Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match:      100.00%     Indels:      0  
DB:              9           Gaps:      0
```

US-10-726-967A-57 (1-7) x AL159987 (1-119024)

QY 1 G|U|U|A|S|n|l|e|n|g|t|u|r|t|h|r|a|p 7

DB 15015 GAATCAACCTTGAAACAGAT 15035

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 26, 2005, 01:07:07 ; Search time 364 Seconds
(without alignments)
113.841 Million cell updates/sec

Title: US-10-726-967a-57
Perfect score: 35
Sequence: 1 ENLEPDT 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=rlh
-Q/cgn2.1/USFTO.spool.h/US10726967/runat_25082005.170152.275/app.query.fasta.1.199
-DB=N.Geneseq -QPM=faacap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -MINLEN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10726967 @CCN 1.1 470 @runat_25082005.170152.275 -NCPU=6 -ICPU=3
-NO WMAP -LANG=JUTERY -NEG=SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: N.Geneseq.16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	length	DB ID	Description
C 1	35	100.0	347	2	AAx83414
C 2	35	100.0	347	2	AAx68946
C 3	35	100.0	347	3	AAx80937
C 4	35	100.0	347	6	AAx9783
C 5	35	100.0	347	6	ABK46827

C 6	35	100.0	347	8	ADA11304	Ada11304 Human bre
C 7	35	100.0	347	10	ADCI5277	Adci5277 Human bre
C 8	33	94.3	42	12	ADP83955	Adp83955 Human BAC
C 9	33	94.3	47319	4	AAK72230	Aak72230 Human imm
10	33	94.3	47319	4	AAK64813	Aak64813 Human imm
11	33	94.3	105413	12	AD136512	Adi36512 Human kin
C 12	33	94.3	236303	4	AASt1614	Aas1614 Human gen
13	32	91.4	105	12	ADK71098	Adk71098 C. elegans
14	32	91.4	232	6	ABL88506	Ab188506 C. caracate
15	32	91.4	235	6	ABL88550	Ab188550 C. tesabula
16	32	91.4	238	6	ABL88541	Ab188541 C. quercin
17	32	91.4	238	6	ABL88540	Ab188540 C. pulicar
18	32	91.4	241	6	ABL88489	Ab188489 C. bandanu
19	32	91.4	256	6	ABL88549	Ab188549 C. striatru
20	32	91.4	257	6	ABL88518	Ab188518 C. gloriat
21	32	91.4	262	6	ABL88578	Ab188578 C. stercus
22	32	91.4	275	6	ABL88566	Ab188566 C. lyncus
23	32	91.4	287	6	ABL88558	Ab188558 C. figulin
24	32	91.4	290	6	ABL88592	Ab188592 C. omaria
25	32	91.4	290	6	ABL88599	Ab188599 C. aulicus
26	32	91.4	293	6	ABL88591	Ab188591 C. omaria
27	32	91.4	293	6	ABL88593	Ab188593 C. omaria
28	32	91.4	296	6	ABL88533	Ab188533 C. marmore
29	32	91.4	298	6	ABL88563	Ab188563 C. quercin
30	32	91.4	304	6	ABL88520	Ab188520 C. latercu
31	32	91.4	313	6	ABL88521	Ab188521 C. latercu
32	32	91.4	315	6	ABL88603	Ab188603 C. consors
33	32	91.4	322	6	ABL88499	Ab188499 C. bullatu
34	32	91.4	325	6	ABL88493	Ab188493 C. bullatu
35	32	91.4	326	6	ABL88494	Ab188494 C. bullatu
36	32	91.4	327	6	ABL88485	Abx04885 Comus sp
37	32	91.4	331	6	ABL88495	Ab188495 C. bullatu
38	32	91.4	336	6	ABL88543	Ab188543 C. radiatu
39	32	91.4	337	6	ABL88497	Ab188497 C. bullatu
40	32	91.4	337	6	ABL88496	Ab188496 C. bullatu
41	32	91.4	337	6	ABL88537	Ab188537 C. marmore
42	32	91.4	415	9	ACH45710	ACH45710 Human foe
C 43	32	91.4	426	5	ABV10971	Abv10971 Human pro
C 44	32	91.4	429	8	ABK62059	Abk62059 Arabidops
C 45	32	91.4	446	5	ABV01802	Abv01802 Human pro

ALIGNMENTS

RESULT 1	AAx83414/c
ID	AAx83414 standard; cDNA; 347 BP.
AC	AAx83414;
XX	
DT	31-AUG-1999 (first entry)
XX	
DE	Breast cancer tumour specific clone #168.
XX	
KW	Breast cancer; tumour; gene expression; genome; diagnosis; mamma;
KM	human endogenous retrovirus; vaccine; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO9725426-A2.
XX	
PD	17-JUL-1997.
XX	
PF	10-JAN-1997; 97MO-US0000485.
XX	
PR	11-JAN-1996; 96US-00585392.
PR	20-AUG-1996; 96US-00700014.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Frudakis TN, Smith JM, Reed SG;
XX	
DR	WPI; 1997-372865/34.

XX Breast cancer-related DNA from retrovirus antigen (s) - useful for
 PT diagnosis and treatment of breast cancer.
 XX
 PS Claim 1, Page 186; 221pp; English.
 XX
 CC Sequences AAX83201-X83285 and AAX83331-X83415 represent novel breast
 CC cancer tumour specific clones which are expressed from a genomic region
 CC containing a human endogenous retrovirus (AAX83330). Detection of the
 CC clone sequences allows determination of the presence of breast cancer in
 CC a mammal. Progression of breast cancer can be monitored by detecting the
 CC level of clone expression. Polypeptides encoded by the clones can be used
 CC in vaccines to inhibit or prevent breast cancer
 XX
 SQ Sequence 347 BP; 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 29.5 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AAX83414 (1-347)

QY 1 GIUleAenLeuGIuThrAsp 7
 DB 114 GAAATTAATTAGAAACAGAT 94

RESULT 2
 AAV68946/c
 ID AAV68946 standard; DNA; 347 BP.
 XX
 AC AAV68946;
 XX
 XX 22-JAN-1999 (first entry)
 XX
 DE DNA molecule encoding a breast tumour specific polypeptide #138.
 XX
 KW Human; breast cancer; breast tumour tissue; diagnosis; treatment;
 XX vaccine; epitope; endogenous; retroviral element; ss.
 OS Homo sapiens.
 XX
 PN MO9845328-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 09-APR-1998; 98WO-US006939.
 XX
 PR 09-APR-1997; 97US-00838762.
 XX
 PR 11-DEC-1997; 97US-00991789.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Frudakis TN, Smith JM, Reed SG;
 XX
 DR WPI; 1998-557473/47.
 XX
 PT New DNA sequences isolated from endogenous human retroviral element - and
 PT related vectors, transformed cells, proteins and antibodies, useful for
 PT diagnosis, treatment and prevention of breast cancer.
 XX
 PS Claim 1, Page 111; 173pp; English.
 XX
 CC AAV68800 to AAV68998 represent nucleotide sequences which encode human
 CC breast tumour specific polypeptides. Detection or measurement of human
 CC breast tumour specific polypeptides and nucleotide sequences, or the
 CC corresponding RNA in a sample, is used for diagnosis and monitoring of
 CC breast cancer. Human breast tumour specific polypeptides and nucleotide
 CC sequences, and the vectors containing the DNAs, are also useful in
 CC vaccines for inhibiting development (for prevention or therapy) of breast

CC cancer. The polypeptides may also be used to raise monoclonal antibodies,
 CC used as immunoassay reagents
 XX
 SQ Sequence 347 BP; 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 29.5 Length: 347
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AAV68946 (1-347)

QY 1 GIUleAenLeuGIuThrAsp 7
 DB 114 GAAATTAATTAGAAACAGAT 94

RESULT 3
 AAC80937/c
 ID AAC80937 standard; CDNA; 347 BP.
 XX
 AC AAC80937;
 XX
 XX 13-FEB-2001 (first entry)
 XX
 DE Human breast tumour-specific CDNA SEQ ID NO: 225.
 XX
 KW Human; breast tumour-specific antigen; cytostatic; vaccine;
 XX breast cancer; B18Ag1; B11Ag1; B15Ag1; ss.
 OS Homo sapiens.
 XX
 PN WO200061753-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-2000; 2000WO-US009312.
 XX
 PR 09-APR-1999; 99US-00289198.
 XX
 PR 28-OCT-1999; 99US-00429755.
 XX
 PR 23-MAR-2000; 2000US-00534825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Frudakis TN, Smith JM, Reed SG, Misher LE, Retter MW, Dillon DC;
 XX
 DR WPI; 2000-628403/60.
 XX
 PT An isolated polypeptide comprising an immunogenic portion of a breast
 PT tumor protein used for inhibiting the development of cancer, especially
 PT breast cancer, and monitoring cancer progression in a patient.
 XX
 PS Claim 4, Page 149; 187pp; English.
 XX
 CC The present sequence is given in a specification relating to compositions
 CC and methods for the treatment and diagnosis of breast cancer. Nucleotide
 CC sequences that are preferentially expressed in breast tumour tissue, and
 CC the polypeptides encoded by such nucleotide sequences, are used in
 CC compositions and vaccines to inhibit the development of cancer,
 CC especially breast cancer. The progression of a cancer may be monitored by
 CC carrying out detection of tumour-specific antigens at subsequent time
 CC points and comparing the results from the different time points. CD4+
 CC and/or CD8+ T-Cells isolated from the cancer patient may be treated with
 CC tumour-specific polypeptides, polynucleotides encoding the polypeptides
 CC or antigen presenting cells expressing the polypeptides. The cells are
 CC then administered to the patient to inhibit development of cancer
 XX
 SQ Sequence 347 BP; 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 29.5 Length: 347

Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-726-967A-57 (1-7) x AAC80937 (1-347)

Qy 1 Glu1leasnleuglurhrasp 7
|||||
Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 4
AAS9783/c
ID AAS9783 standard; cDNA; 347 BP.
XX

AC AAS9783;

XX 12-MAR-2002 (first entry)

XX Breast tumour-specific DNA #158.

XX Human; breast cancer; PCR primer; ss; cytostatic; immunostimulant;

XX tumour; vaccine; immunogenic.

XX Homo sapiens.

XX WO200190152-A2.

XX 29-NOV-2001.

XX 22-MAY-2001; 2001WO-US016776.

XX 24-MAY-2000; 2000US-00577505.

PR 06-JUN-2000; 2000US-00590583.

PR 26-OCT-2000; 2000US-0069295.

PR 16-MAR-2001; 2001US-00810936.

XX (CORI-) CORIXA CORP.

XX Frudakis TN, Reed SG, Smith JM, Misher LE, Dillon DC, Retter MW,

PI Wang A, Skeiky YAM, Harlocker SL, Day CH;

XX WPI; 2002-089919/12.

XX New breast tumor proteins and polynucleotides encoding them, useful for

PT treating and/or preventing cancer, particularly breast cancer, and for

XX eliciting humoral and/or cellular immune response.

PS Claim 1; Page 197; 245pp; English.

XX The invention relates to novel breast tumour polynucleotides and

CC polypeptides. The polypeptides and polynucleotides are useful in

CC pharmaceutical compositions for treating and/or preventing cancer,

CC particularly breast cancer, and for eliciting an immune response,

CC may be used as probes or primers for nucleic acid hybridisation, in the

CC design and preparation of ribozyme molecules for inhibiting expression of

CC tumour polypeptides and proteins, and in recombinant DNA molecules to

CC direct expression of a polypeptide in host cells. AAS9570-AAS9888

CC represent novel human breast cancer protein coding sequences and PCR

XX primers of the invention

XX

SO Sequence 347 BP; 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 29.5

XX Score: 35.00

XX Percent Similarity: 100.00%

XX Best Local Similarity: 100.00%

XX Query Match: 100.00%

US-10-726-967A-57 (1-7) x AAS9783 (1-347)

Qy 1 Glu1leasnleuglurhrasp 7
|||||
Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 5
ABK46827/c
ID ABK46827 standard; cDNA; 347 BP.
XX

AC ABK46827;

XX 05-JUN-2002 (first entry)

XX Human breast tumour-specific cDNA #145.

XX Human; breast tumour-specific protein; vaccine; breast cancer; gene; ss.

XX Homo sapiens.

XX US6344550-B1.

XX 05-FEB-2002.

XX 17-APR-1998; 98US-00062451.

XX 01-JAN-1996; 96US-00585392.

PR 20-AUG-1996; 96US-00700014.

PR 10-JAN-1997; 97MO-US000485.

PR 09-APR-1997; 97US-0083762.

PR 11-DEC-1997; 97US-00991789.

XX (CORI-) CORIXA CORP.

XX Frudakis TN, Smith JM, Reed SG;

PI WPI; 2002-215084/27.

XX Polynucleotide encoding breast-specific tumor polypeptides useful as

PT vaccine for preventing and treating breast cancer in a subject.

XX Example 1; Col 159; 128pp; English.

XX The invention relates to an isolated DNA molecule (I) encoding breast-

CC tumour-specific polypeptides. (I) is useful as a vaccine for preventing

CC and treating breast cancer in a subject. The polypeptide encoded by (I)

CC is used for production of compounds such as antibodies useful in

CC diagnosing and monitoring the progression of breast cancer. ABK4614-

CC ABK4689 represent human breast tumour-specific coding sequences and

CC related PCR primers of the invention

XX

SO Sequence 347 BP; 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 29.5

XX Score: 35.00

XX Percent Similarity: 100.00%

XX Best Local Similarity: 100.00%

XX Query Match: 100.00%

XX DB: 6 Gaps: 0

US-10-726-967A-57 (1-7) x ABK46827 (1-347)

Qy 1 Glu1leasnleuglurhrasp 7
|||||
Db 114 GAAATTAAATTAGAAACAGAT 94

RESULT 6
ADA11304/c
ID ADA11304 standard; cDNA; 347 BP.
XX

XX ADA11304;

XX

CC breast cancer comprises administering a composition comprising breast
CC tumour proteins and nucleic acids, which stimulates and/or expands T cells
CC specific for the tumour protein. The present sequence was used to
CC illustrate the invention.

XX
SQ Sequence 347 BP, 102 A; 87 C; 68 G; 90 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	29, 5	Length:	347
Score:	35.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-726-967a-57 (1-7) x ADC15277 (1-347)

QY 1 GluIleAsnLeuGluThrAsp 7
DB 114 GAATTAATTAGAAACAGAT 94

RESULT 8

ADP83955/C
ID ADP83955 standard; DNA; 42 BP.

XX
AC ADP83955;

XX
DT 23-SEP-2004 (first entry)

DE Human BACE1 mutagenic oligonucleotide SEQ ID NO:82.

XX human; beta-site amyloid precursor protein cleaving enzyme 1;

KM beta-site APP cleaving enzyme 1; BACE1; BACE1 isoform A; chromosome 11;

KM prodomain; engineered cleavage site; protease domain; neuroprotective;

KM neurotrophic; gene therapy; Alzheimer's disease; Down's syndrome;

KM mutagenesis; ss.

XX

OS Homo sapiens.

OS Synthetic.

XX
PN WO2004056962-A2.

XX
PD 08-JUL-2004.

XX
PF 02-DEC-2003; 2003WO-US038314.

XX
PR 04-DEC-2002; 2002US-0430984P.

XX
PA (SUNE-) SUNESIS PHARM INC.

XX
PI Ballinger M;

XX
DR WPI; 2004-507703/48.

XX
PT New polypeptides for producing homogenously processed preparations of

PT beta site amyloid precursor protein-cleaving enzyme comprises a

PT prodomain, an engineered cleavage site and a protease domain.

XX
PS Example 1; SEQ ID NO 82; 40pp; English.

XX
CC The present invention describes a polypeptide (I) comprising in order

CC from the N-terminus to the C-terminus: (a) a prodomain comprising at

CC least 6 contiguous amino acids of the 16 amino acid sequence of SEQ ID

CC NO:3 (ADP83877), comprising residues 22-37 of SEQ ID NO:1 (ADP83876) which

CC is the longest isoform of human beta-site amyloid precursor protein (APP)

CC cleaving enzyme 1 (BACE1), isoform A; (b) an engineered cleavage site;

CC and (c) a protease domain. (I) is capable of being cleaved at the

CC engineered cleavage site, and so releases a free protease domain that has

CC BACE1 activity. Also described: (1) a nucleic acid sequence encoding (I);

CC (2) a vector for expression of (I); and (3) a host cell expressing (I);

CC (I) has neuroprotective and neurotrophic activities, and can be used in gene

CC therapy. (I) can be used for producing preparations of homogenously

CC processed BACE that may be used for e.g. studying or treating diseases

CC such as Alzheimer's disease or Down's syndrome. The human BACE1 gene is

CC located on chromosome 11, more specifically to 11q23.2-23.3. The present

CC sequence represents a human BACE1 mutagenic oligonucleotide, which is

CC used in an example from the present invention.

XX
SQ Sequence 42 BP, 6 A; 14 C; 11 G; 11 T; 0 U; 0 Other;

XX

US-10-726-967a-57 (1-7) x ADP83955 (1-42)

QY 1 GluIleAsnLeuGluThrAsp 7
DB 27 GAACGAACTCGAGACCGAC 7

RESULT 9

AAK72230
ID AAK72230 standard; DNA; 47319 BP.

XX
AC AAK72230;

XX
DT 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27042.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KM cytostatic; gene therapy; vaccine; metastasis; ds.

XX

OS Homo sapiens.

XX
PN WO200157182-A2.

XX
PD 09-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US001354.

XX
PR 31-JAN-2000; 2000US-0179065P.

XX
PR 04-FEB-2000; 2000US-0180628P.

XX
PR 24-FEB-2000; 2000US-0184664P.

XX
PR 02-MAR-2000; 2000US-0186350P.

XX
PR 16-MAR-2000; 2000US-0189874P.

XX
PR 17-MAR-2000; 2000US-0190076P.

XX
PR 18-APR-2000; 2000US-0198123P.

XX
PR 19-MAY-2000; 2000US-0205515P.

XX
PR 07-JUN-2000; 2000US-0209467P.

XX
PR 28-JUN-2000; 2000US-0214886P.

XX
PR 30-JUN-2000; 2000US-0215135P.

XX
PR 07-JUL-2000; 2000US-0216647P.

XX
PR 11-JUL-2000; 2000US-0217487P.

XX
PR 11-JUL-2000; 2000US-0217496P.

XX
PR 14-JUL-2000; 2000US-0218290P.

XX
PR 26-JUL-2000; 2000US-0220963P.

XX
PR 26-JUL-2000; 2000US-0220964P.

XX
PR 14-AUG-2000; 2000US-0224518P.

XX
PR 14-AUG-2000; 2000US-0224519P.

XX
PR 14-AUG-2000; 2000US-0225213P.

XX
PR 14-AUG-2000; 2000US-0225214P.

XX
PR 14-AUG-2000; 2000US-0225266P.

XX
PR 14-AUG-2000; 2000US-0225267P.

XX
PR 14-AUG-2000; 2000US-0225268P.

```

PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236386P.
PR 29-SEP-2000; 2000US-0236396P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.

PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251038P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 27042; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK5950 and AAM82169
XX represent sequences used in the exemplification of the present invention.

SQ Sequence 47319 BP; 13708 A; 8453 C; 9720 G; 15438 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.04e+04 Length: 47319
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-726-967A-57 (1-7) x AAK72230 (1-47319)

```

Qy 1 GluileasnlenglutrRasp 7
|||::|||
Db 15242 GAGTTAACTTGAGACAGAT 15262

RESULT 10
AAK64813
ID AAK64813 standard; DNA; 47319 BP.
XX
AC AAK64813;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19625.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
OS Cytotoxic; gene therapy; vaccine; metastasis; db.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
PD
XX 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216860P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232402P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237032P.
PR 02-OCT-2000; 2000US-0237033P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

KW	transgenic; drug screening assay; gene therapy; chromosome X; gene; ds
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	2388..102412
FT	CD5
FT	/*tag= b
FT	/product= "Human kinase protein"
FT	2388..2553
FT	/*tag= a
FT	2554..39832
FT	/*tag= c
FT	variation
FT	replace(2916..A)
FT	/*tag= d
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(3308..G)
FT	/*tag= e
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(4640..G)
FT	/*tag= f
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(4674..A)
FT	/*tag= g
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(5255..G)
FT	/*tag= h
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6268..C)
FT	/*tag= i
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6437..6439..CT)
FT	/*tag= j
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6456..6458..TG)
FT	/*tag= k
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6486..C)
FT	/*tag= l
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6534..T)
FT	/*tag= m
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(6595..G)
FT	/*tag= n
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(9622..9624..GT)
FT	/*tag= o
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(13172..T)
FT	/*tag= p
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(14802..C)
FT	/*tag= q
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(20393..T)
FT	/*tag= r
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(22253..A)
FT	/*tag= s
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(22253..T)
FT	/*tag= t
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(25890..A)
FT	/*tag= u
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(25891..A)
FT	/*tag= v
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(25894..C)
FT	/*tag= w
FT	/standard_name= "Single nucleotide polymorphism (SNP)"
FT	replace(27386..C)
FT	/*tag= x
FT	/standard_name= "Single nucleotide polymorphism (SNP)"


```

FT      /*tag= x
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(28860..28862,Tx)
FT      /*tag= y
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(25861,T)
FT      /*tag= z
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(30916,G)
FT      /*tag= aa
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(30959,A)
FT      /*tag= ab
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(34434..34436,Tx)
FT      /*tag= ac
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(36813,T)
FT      /*tag= ad
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(37063,T)
FT      /*tag= ae
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(37569,C)
FT      /*tag= af
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(37905,A)
FT      /*tag= ag
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      39833..40001
FT      /*tag= ah
FT      40002..59181
FT      /*tag= ai
FT      replace(41617,T)
FT      /*tag= aj
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(42444,T)
FT      /*tag= ak
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(46230,A)
FT      /*tag= al
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(46604,T, A)
FT      /*tag= am
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(46867,C)
FT      /*tag= an
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(51123,G)
FT      /*tag= ao
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(51212,G)
FT      /*tag= ap
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(52308,C)
FT      /*tag= aq
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(52333,T)
FT      /*tag= ar
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(53079,C)
FT      /*tag= as
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      59182..59445
FT      /*tag= at
FT      59446..88092
FT      /*tag= au
FT      replace(62141,T)
FT      /*tag= av
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(70338..70340,CT)
FT      /*tag= aw
FT      /standard_name= "Single nucleotide polymorphism (SNP) "

```

```

FT      variation
FT      replace(70379,A)
FT      /*tag= ax
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(76074,T)
FT      /*tag= ay
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(77336,C)
FT      /*tag= az
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(77582,C)
FT      /*tag= ba
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(78010,C, T)
FT      /*tag= bb
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(78011,G, C)
FT      /*tag= bc
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(79557,G)
FT      /*tag= bd
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(79617,C)
FT      /*tag= be
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(84900,G)
FT      /*tag= bf
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      88093..88188
FT      /*tag= bg
FT      88189..93291
FT      /*tag= bh
FT      93292..93349
FT      /*tag= bi
FT      93350..98719
FT      /*tag= bj
FT      98720..98797
FT      /*tag= bk
FT      98798..102286
FT      /*tag= bl
FT      replace(100529,G)
FT      /*tag= bm
FT      /standard_name= "Single nucleotide polymorphism (SNP) "
FT      replace(100996..100998,CC)
FT      /*tag= bn
FT      /standard_name= "Single nucleotide polymorphism (SNP) "

```

Alignment Scores:

```

Pred. No.: 4.94e+04 Length: 105413
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 12 Gaps: 0

```

US-10-726-967A-57 (1-7) x ADI36512 (1-105413)

```

Qy      1 GluIIeASnleuGIurhPrp 7
Db      45277 GAGATCAACATTGAGACAGAT 45297

```

RESULT 12

AA511614/c
ID AA511614 standard; DNA; 236303 BP.

XX AA511614;

DT 24-OCT-2001 (first entry)

DE Human genomic DNA containing exons 2-17 of the CRIM1 gene.

KW CRIM-1; Human; human chromosome 2p21-16.3; ophthalmological;
KW neuroprotective; renal; osteopathic; dental; vulnerrary; immunogen;
KW antibody; gene therapy; neurodegenerative disease; eye disorder;

KW cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
 KW tooth abnormality; wound; de.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 33104..33277
 FT /tag= a
 FT /number= 2
 FT intron 33278..77746
 FT /tag= b
 FT /number= 2
 FT /tag= c
 FT /number= 3
 FT intron 77990..79103
 FT /tag= d
 FT /number= 3
 FT exon 79104..79224
 FT /tag= e
 FT /number= 4
 FT intron 79225..101022
 FT /tag= f
 FT /number= 4
 FT exon 101023..101144
 FT /tag= g
 FT /number= 5
 FT intron 101145..113377
 FT /tag= h
 FT /number= 5
 FT exon 113378..113560
 FT /tag= i
 FT /number= 6
 FT intron 113561..115985
 FT /tag= j
 FT /number= 6
 FT exon 115986..116183
 FT /tag= k
 FT /number= 7
 FT intron 116184..135707
 FT /tag= l
 FT /number= 7
 FT exon 135708..135836
 FT /tag= m
 FT /number= 8
 FT intron 135837..146471
 FT /tag= n
 FT /number= 8
 FT exon 146472..146628
 FT /tag= o
 FT /number= 9
 FT intron 146629..148761
 FT /tag= p
 FT /number= 9
 FT exon 148762..148883
 FT /tag= q
 FT /number= 10
 FT intron 148884..150044
 FT /tag= r
 FT /number= 10
 FT exon 150045..150254
 FT /tag= s
 FT /number= 11
 FT intron 150255..153815
 FT /tag= t
 FT /number= 11
 FT exon 153816..154031
 FT /tag= u
 FT /number= 12
 FT intron 154032..158580
 FT /tag= v
 FT /number= 12
 FT exon 158581..158802

FT /tag= w
 FT /number= 13
 FT intron 158803..173982
 FT /tag= x
 FT /number= 13
 FT exon 173983..174177
 FT /tag= y
 FT /number= 14
 FT intron 174178..181006
 FT /tag= z
 FT /number= 14
 FT exon 181007..181129
 FT /tag= aa
 FT /number= 15
 FT intron 181130..183612
 FT /tag= ab
 FT /number= 15
 FT exon 183613..183800
 FT /tag= ac
 FT /number= 16
 FT intron 183801..185152
 FT /tag= ad
 FT /number= 16
 FT exon 185153..187765
 FT /tag= ae
 FT /number= 17
 XX WO200138519-A1.
 XX 31-MAY-2001.
 XX 24-NOV-2000; 2000WO-AU001435.
 XX 26-NOV-1999; 99AU-00004348.
 XX (UYOU) UNIV QUEENSLAND.
 XX Little M, Yamada T, Holmes G, Georgae K, Kolle G, Wilkinson L,
 XX WPI; 2001-343951/36.
 DR
 XX
 FT Nucleic acids from human chromosome 2p21-16.3 and the encoded peptide,
 PT useful for preventing, diagnosing and treating e.g. eye disease,
 FT especially cataract formation.
 PS
 XX Claim 4; Fig 3; 16pp; English.
 XX
 CC The invention relates to nucleic acids from human chromosome 2p21-16.3
 CC and the encoded peptide (and mouse and chicken orthologues) that
 CC comprises a pGECCLP group, an insulin-like growth factor binding protein
 CC (IGFBP)-like domain, cysteine-rich domains, an RGD (undefined) group and
 CC a transmembrane domain. The protein, e.g. CRIM1, interacts with peptides
 CC of the transforming growth factor superfamily. A composition comprising
 CC an expression construct comprising the nucleic acids of the invention or
 CC a mimetic which antagonises or mimics an activity of a CRIM1 polypeptide
 CC may be used in a method for modulating the biological activity of a
 CC polypeptide of the bone morphogenic protein (BMP) family. In this way
 CC they may be used to prevent or treat an eye disease, especially cataract
 CC formation. They may also be used to treat neurodegenerative diseases,
 CC renal and kidney disease, bone and tooth abnormalities, wounds and skin
 CC damage, e.g. by use of the nucleic acid in gene therapy by using
 CC antibodies directed against CRIM1 polypeptides. The present sequence is a
 CC Human genomic DNA containing exons 2-17 of the CRIM1 gene
 XX
 SQ Sequence 236303 BP; 70199 A; 46544 C; 47996 G; 71563 T; 0 U; 1 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.2e+05 Length: 236303
 Score: 33.00 Matches: 6
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 85.71% Mismatches: 0
 Query Match: 94.29% Indels: 0
 DB: 4 Gaps: 0

US-10-726-967A-57 (1-7) x AAS11614 (1-236303)

Qy 1 Glu1leAenleuGIuThraap 7

Db 212401 GAGTTAACTTGAGACGAGAT 212381

RESULT 13

ADK71098
ID ADK71098 standard; DNA; 105 BP.

AC ADK71098;

DT 06-MAY-2004 (first entry)

DE C. elegans SL-1 exon/Intron fragment, SEQ ID 3.

KM Chimeric ribonucleic acid; trans-splicing nucleic acid; gene analysis;
KW transcripction mapping; SL-1; ds.

XX Caenorhabditis elegans.

OS
PN W02004015085-A2.

PD 19-FEB-2004.

PF 11-AUG-2003; 2003WO-US025081.

PR 09-AUG-2002; 2002US-0402473P.

PR 04-NOV-2002; 2002US-0423490P.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

PI Sternberg P, Hwang BJ;

XX WPI; 2004-169499/16.

PT Producing 5'-labeled chimeric cDNA from a cDNA preparation from cells expressing a 5' trans-splicing nucleic acid. The
PT expressing a 5'-trans-splicing nucleic acid, useful for identifying and
PT analyzing genes and RNA transcripts in samples.

XX Disclosure; Page 28; 113pp; English.

CC The invention relates to producing 5'-labeled chimeric cDNA from a cDNA
CC preparation from cells expressing a 5' trans-splicing nucleic acid. The
CC method involves forming a mixture comprising: a cDNA preparation derived
CC from cells expressing a 5'-trans-splicing nucleic acid comprising an exon
CC and an intron; a labelled oligonucleotide comprising a sequence that
CC hybridizes to at least a portion of the exon of the 5'-trans-splicing
CC nucleic acid; and an enzyme that catalyzes polynucleotide synthesis; and
CC incubating the mixture under conditions that permit polynucleotide
CC synthesis. The methods and nucleic acid constructs are useful for
CC producing 5'-labeled chimeric RNAs and cDNAs. The 5'-labeled chimeric
CC RNAs and cDNAs can be used for identifying and analyzing genes and RNA
CC transcripts in samples, mapping transcription start sites in genomes,
CC identifying novel transcription start sites, characterizing global gene
CC and protein expression in specific cell types, or preparing probe arrays
CC that hybridize to sequences from the 5'-regions of transcript. The
CC present sequence represents a C. elegans SL-1 exon/Intron fragment, an
CC example of 5'-trans-splicing sequence.

XX SQ Sequence 105 BP; 37 A; 18 C; 21 G; 29 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 40.7 Length: 105
Score: 32.00 Matches: 5
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 71.43% Mismatches: 0
Query Match: 91.43% Indels: 0
DB: 12 Gaps: 0

US-10-726-967A-57 (1-7) x ADK71098 (1-105)

Qy 1 Glu1leAenleuGIuThraap 7

Db 20 GAGTTAACTTGAGAACTGAC 40

RESULT 14

ABL8506
ID ABL8506 standard; DNA; 232 BP.

AC ABL8506;

DT 24-MAY-2002 (first entry)

DE C characteristic mu-conopeptide Ca3-6 coding sequence.

XX Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;
XX anaesthetic; neuromuscular blocking agent; neuroprotective; pain;
KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;
KW cardiovascular; vasotropic; cardiant; tranquilizer; antimigraine;
KW neurodegenerative disease; neuromuscular disorder; gene; ds.

XX Conus characteristic.

OS
PN W0200207678-A2.

PD 31-JAN-2002.

PF 23-JUL-2001; 2001WO-US023125.

PR 21-JUL-2000; 2000US-0219619P.

PR 03-NOV-2000; 2000US-0245157P.

PR 21-MAR-2001; 2001US-0264319P.

XX (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Olivera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;

PI Jacobsen R, Jones RM, Cartier GE, Shen GS;

XX WPI; 2002-217020/27.

DR P-PSDB; ABB88187.

PT New mu-conopeptides useful for treating disorders associated with voltage
PT -gated sodium channels, e.g. stroke or pain, as neuromuscular blocking
PT agents, as local anesthetic agents, as analgesic agents and as
PT neuroprotective agents.

XX Claim 8; Page 34; 231pp; English.

CC The present invention relates to mu-conopeptides derived from snails,
CC which can be in the treatment of disorders associated with voltage-gated
CC ion channels. These may include neurodegenerative disorders such as
CC amyotrophic lateral sclerosis, neuromuscular disorders, post-operative or
CC severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,
CC post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,
CC epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia
CC or ischaemia which typically follows stroke, cerebrovascular accident,
CC brain or spinal cord trauma, myocardial infarct, physical trauma,
CC drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for
CC providing muscle relaxation, treating essential blepharospasm and other
CC forms of focal dystonia, and for anti-wrinkle use. The present sequence
CC is a mu-conopeptide preproteins coding sequence of the invention

XX SQ Sequence 232 BP; 44 A; 61 C; 56 G; 71 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 97.9 Length: 232
Score: 32.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 91.43% Indels: 0
DB: 6 Gaps: 0

US-10-726-967a-57 (1-7) x ABL88506 (1-232)

Qy 1 Glu1eAsnLeuGluThrAsp 7
 |||||
 Db 80 GAGATCAACCTGCGACGCCGAC 100

RESULT 15

ABL88550
 ID ABL88550 standard; DNA; 235 BP.

AC ABL88550;

XX 24-MAY-2002 (first entry)

DE C tessulatus mu-conopeptide Tε3-1 coding sequence.

XX Mu-conopeptide; snail; venom; voltage-gated sodium channel; analgesic;
 KW anaesthetic; neuromuscular blocking agent; neuroprotective; pain;
 KW cerebroprotective; anticonvulsant; antiaging; antidiabetic;
 KW cardiovascular; vasotropic; cardiant; tranquilizer; antimigraine;
 KW neurodegenerative disease; neuromuscular disorder; gene; ds.

XX Conus tessulatus.

OS MO200207678-A2.

XX 31-JAN-2002.

XX 23-JUL-2001; 2001WO-US023125.

XX 21-JUL-2000; 2000US-0219619P.

PR 03-NOV-2000; 2000US-0245157P.

PR 29-JAN-2001; 2001US-0245157P.

PR 21-MAR-2001; 2001US-0277270P.

XX (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Oliviera BM, McIntosh JM, Garrett JE, Watkins M, Cruz LJ, Shon K;

PI Jacobsen R, Jones RM, Cartier GE, Shen GS;

XX WPI; 2002-217020/27.

DR P-PSDB; ABB88285.

XX New mu-conopeptides useful for treating disorders associated with voltage

PT gated sodium channels, e.g. stroke or pain, as neuromuscular blocking

PT agents, as local anesthetic agents, as analgesic agents and as

PT neuroprotective agents.

XX Claim 8; Page 56; 231pp; English.

XX The present invention relates to mu-conopeptides derived from snails,

XX which can be in the treatment of disorders associated with voltage-gated

XX ion channels. These may include neurodegenerative disorders such as

XX amyotrophic lateral sclerosis, neuromuscular disorders, post-operative or

XX severe chronic pain, stroke, trigeminal neuralgia, diabetic neuropathy,

XX post-herpetic neuralgia, neuronal pain and phantom limb, burn pain,

XX epilepsy, for reducing neurotoxic injury associated with hypoxia, anoxia

XX or ischemia which typically follows stroke, cerebrovascular accident,

XX brain or spinal cord trauma, myocardial infarct, physical trauma,

XX drowning, suffocation, perinatal asphyxia, or hypoglycaemic events, for

XX providing muscle relaxation, treating essential blepharospasm and other

XX forms of focal dystonia, and for anti-wrinkle use. The present sequence

XX is a mu-conopeptide preprotein coding sequence of the invention

XX SQ Sequence 235 BP; 50 A; 58 C; 57 G; 70 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 99.3

XX Score: 32.00

XX Percent Similarity: 100.00%

XX Best Local Similarity: 85.71%

XX Query Match: 91.43%

XX Length: 235

XX Matches: 6

XX Conservative: 1

XX Mismatches: 0

XX Indels: 0

DB: 6 Gaps: 0

US-10-726-967a-57 (1-7) x ABL88550 (1-235)

Qy 1 Glu1eAsnLeuGluThrAsp 7
 |||||
 Db 80 GAGATCAACCTGCGACGCCGAC 100

Search completed: August 26, 2005, 01:17:55
 Job time : 390 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 26, 2005, 01:09:23 ; Search time 2143 Seconds
(without alignments)
124.335 Million cell updates/sec

Title: US-10-726-967a-57
Perfect score: 35
Sequence: 1 EINFELD 7

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O/cgnt 1/USPTO.spool_h/US10726967/runat_25082005_170153_293/app.query.fasta_1.199
-DB=EST -QPM=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-OUTCAL=GN-200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-USPR=US10726967 @CGN 1.1 3437 @runat_25082005_170153_293 -NCPU=6
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DELTIMEOUT=120 -WANT TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	250	4	BG046764
2	35	100.0	431	2	AV961777
3	35	100.0	485	1	AV947666
4	35	100.0	526	2	AV984343
5	35	100.0	527	2	AV963752
6	35	100.0	528	9	CE258195
7	35	100.0	541	5	BM037056
8	35	100.0	567	2	AV982887
9	35	100.0	579	5	BM212514

10	35	100.0	612	5	BM192913
11	35	100.0	618	5	BM220061
12	35	100.0	628	2	AV959453
13	35	100.0	638	5	BM220124
14	35	100.0	657	5	BM208590
15	35	100.0	684	5	BM239812
16	35	100.0	685	5	BM220246
17	35	100.0	695	5	BM274369
18	35	100.0	716	5	BM240676
19	35	100.0	723	5	BM248343
20	35	100.0	925	6	CD327193
21	35	100.0	1127	1	AL523587
22	35	100.0	1536	9	CL976221
23	34	97.1	387	6	CB805633
24	34	97.1	420	6	CD472645
25	34	97.1	645	4	B1687905
26	34	97.1	892	3	CNS0ADNM
27	33	94.3	446	8	AZ009511
28	33	94.3	503	7	CN888811
29	33	94.3	522	4	BG515742
30	33	94.3	536	2	BE192331
31	33	94.3	555	8	AZ979045
32	33	94.3	607	9	DR35017S
33	33	94.3	651	7	CN983221
34	33	94.3	653	9	AG041283
35	33	94.3	660	4	BI270807
36	33	94.3	671	7	CN9882754
37	33	94.3	700	9	CE853090
38	33	94.3	701	5	BU725347
39	33	94.3	718	4	BU091620
40	33	94.3	733	8	BH334365
41	33	94.3	749	8	BZ331318
42	33	94.3	756	9	CE377791
43	33	94.3	912	9	CR191302
44	33	94.3	1015	9	AG378872
45	33	94.3	1347	4	B1524484

ALIGNMENTS

RESULT 1
BG046764 250 bp mRNA linear EST 23-JUN-2004
aa60f09.y1 Gm-c1060 Glycine soja cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1060-1050.5, similar to TR-0957L9 Q987L9 SUBUNIT 68 OF
CYTOCHROME C OXIDASE. ;, mRNA sequence.

ACCESSION
BG046764
VERSION
BG046764.1 GI:12495844

KEYWORDS
SOURCE
ORGANISM

Glycine soja
Glycine soja
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
AUTHORS

1 (bases 1 to 250)
Shoemaker, R., Keim, P., Vodka, L., Erpelting, J., Corvett, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.

TITLE
JOURNAL

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. Trace

considered overall poor quality this clone is available through:
 Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone:
 800 423 4163; email: info@biogeneticservices.com)
 High quality sequence scrop: 1.

FEATURES

Location/Qualifiers
 1..250

/organism="Glycine soja"
 /mol_type="mRNA"
 /db_xref="taxon:3848"
 /clone="GENOME SYSTEMS CLONE ID: Gm-C1060-1050"
 /tissue_type="Root, 2 week seedlings"
 /lab_host="DH10B"
 /clone_lib="Gm-C1060"

/note="vector: Bluescript II SK+, Site 1: EcoRI, Site 2: XhoI, The cDNA library was constructed from mRNA isolated from root tissue of 2 week old seedlings for p1468916. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa State University."

ORIGIN

Alignment Scores:

Pred. No.: 124 Length: 250
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-726-967a-57 (1-7) x BG046764 (1-250)

Qy 1 G|U|I|A|S|n|L|e|u|G|U|T|H|A|S|P 7
 |||||
 Db 77 GAAATTAACTCTGAAACTGAT 97

RESULT 2

AV961777

LOCUS AV961777 Nori Satoh unpublished cDNA library, cleavage stage embryo 431 bp mRNA linear EST 14-MAR-2002

DEFINITION Ciona intestinalis cDNA clone c1c120K19 5', mRNA sequence.

ACCESSION AV961777

VERSION AV961777.1 GI:19450076

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

REFERENCE 1 (bases 1 to 431) Phlebobranchia: Clonidae; Ciona.

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoch@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers
 1..431

/organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="c1c120K19"
 /tissue_type="whole animal"
 /dev_stage="cleavage stage embryo"
 /clone_lib="Nori Satoh unpublished cDNA library, cleavage

ORIGIN stage embryo"

Alignment Scores:

Pred. No.: 226 Length: 431
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AV961777 (1-431)

Qy 1 G|U|I|A|S|n|L|e|u|G|U|T|H|A|S|P 7
 |||||
 Db 197 GAAATTAACTCTGAAACTGAT 217

RESULT 3

AV947666

LOCUS AV947666 Nori Satoh unpublished cDNA library, young adult Ciona 485 bp mRNA linear EST 14-MAR-2002

DEFINITION Ciona intestinalis cDNA clone c1ad01101 5', mRNA sequence.

ACCESSION AV947666

VERSION AV947666.1 GI:19425425

KEYWORDS EST.

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

REFERENCE 1 (bases 1 to 485) Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;

AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.

TITLE Expressed genes in Ciona intestinalis

JOURNAL Unpublished (2000)

COMMENT Contact: Nori Satoh

Department of Zoology

Kyoto University

Sakyo-ku, Kyoto, Kyoto 606-8502, Japan

Tel: 81-75-753-4081

Fax: 81-75-705-1113

Email: satoch@ascidian.zool.kyoto-u.ac.jp.

Location/Qualifiers
 1..485

/organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="c1ad01101"
 /tissue_type="whole animal"
 /dev_stage="young adult"
 /clone_lib="Nori Satoh unpublished cDNA library, young

ORIGIN

adult"

Alignment Scores:

Pred. No.: 257 Length: 485
 Score: 35.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-10-726-967a-57 (1-7) x AV947666 (1-485)

Qy 1 G|U|I|A|S|n|L|e|u|G|U|T|H|A|S|P 7
 |||||
 Db 324 GAAATTAACTCTGAAACTGAT 344

RESULT 4

AV984343

LOCUS AV984343 Nori Satoh unpublished cDNA library, egg Ciona 526 bp mRNA linear EST 15-MAR-2002

DEFINITION Ciona intestinalis cDNA clone c1eg29n12 5', mRNA sequence.

ACCESSION AV984343

VERSION AV984343.1 GI:19483134

```

KEYWORDS      EST.
SOURCE         Ciona intestinalis
ORGANISM       Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE      1 (bases 1 to 526)
AUTHORS        Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE          Expressed genes in Ciona intestinalis
JOURNAL        Unpublished (2000)
COMMENT        Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@acidian.zool.kyoto-u.ac.jp.

FEATURES
  source       Location/Qualifiers
               1..526
               /organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /cfeature="cleg29n12"
               /translation="whole animal"
               /dev_stage="egg"
               /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
Pred. No.:      282      Length:      526
Score:          35.00    Matches:      7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%  Indels:      0
DB:             Gaps:      0

US-10-726-967a-57 (1-7) x AV963752 (1-526)

QY      1 G|u|l|e|a|n|l|e|u|g|u|t|u|r|a|s|p 7
Db      241 GAAATTAACCTTGAAACTGAT 261

RESULT 5      527 bp      mRNA      linear      EST 14-MAR-2002
LOCUS       AV963752 Nori Satoh unpublished cDNA library, egg Ciona
DEFINITION  AV963752 Nori Satoh unpublished cDNA library, egg Ciona
ACCESSION   AV963752
VERSION     AV963752.1 GI:19452051
KEYWORDS    EST.
SOURCE      Ciona intestinalis
ORGANISM    Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE    1 (bases 1 to 527)
AUTHORS      Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE        Expressed genes in Ciona intestinalis
JOURNAL      Unpublished (2000)
COMMENT      Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@acidian.zool.kyoto-u.ac.jp.

FEATURES
  source       Location/Qualifiers
               1..527
               /organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /cfeature="cleg15a02"
               /translation="whole animal"
               /dev_stage="egg"
               /clone_lib="Nori Satoh unpublished cDNA library, egg"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      282      Length:      527
Score:          35.00    Matches:      7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%  Indels:      0
DB:             Gaps:      0

US-10-726-967a-57 (1-7) x AV963752 (1-527)

QY      1 G|u|l|e|a|n|l|e|u|g|u|t|u|r|a|s|p 7
Db      238 GAAATTAACCTTGAAACTGAT 258

RESULT 6      528 bp      DNA      linear      GSS 26-SEP-2003
LOCUS       CE258195 tigr-gss-dog-17000345953229 Dog Library Canis familiaris genomic,
DEFINITION  genomic survey sequence.
ACCESSION   CE258195
VERSION     CE258195.1 GI:35964205
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Carnivora; Fissipedidae; Canidae; Canis.
REFERENCE    1 (bases 1 to 528)
AUTHORS      Kirkness,E.F., Batha,V., Halpern,A.L., Levy,S., Remington,K.,
               Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
               Venter,J.C.
TITLE        The dog genome: survey sequencing and comparative analysis
JOURNAL      Science 301 (5641), 1898-1903 (2003)
MEDLINE      22875432
PUBMED       14512627
COMMENT      Contact: Kirkness EF
               The Institute for Genomic Research
               Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
               Rockville, MD 20850, USA
               Tel: 301-838-0200
               Fax: 301-838-0208
               Email: ekirknes@tigr.org
               Class: shotgun.

FEATURES
  source       Location/Qualifiers
               1..528
               /organism="Canis familiaris"
               /mol_type="genomic DNA"
               /strain="Standard Poodle"
               /db_xref="taxon:9615"
               /clone_lib="Dog Library"
               /note="Site 1: B2CX1; Libraries were prepared from
               peripheral blood"

ORIGIN
Alignment Scores:
Pred. No.:      283      Length:      528
Score:          35.00    Matches:      7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%  Indels:      0
DB:             Gaps:      0

US-10-726-967a-57 (1-7) x CE258195 (1-528)

QY      1 G|u|l|e|a|n|l|e|u|g|u|t|u|r|a|s|p 7
Db      423 GAAATCAACCTTGAGACTGAC 443

RESULT 7      541 bp      mRNA      linear      EST 13-OCT-2002
LOCUS       BW037056 Nori Satoh unpublished cDNA library, blood cells Ciona
DEFINITION  BW037056 Nori Satoh unpublished cDNA library, blood cells Ciona

```

```

ACCESSION      intestinalis cDNA clone cibd032115 5', mRNA sequence.
VERSION        BM037056
KEYWORDS       GI:23953008
SOURCE         EST.
ORGANISM       Ciona intestinalis
               Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE      1 (bases 1 to 541)
AUTHORS        Satoh Y., Satake M., Azumi K., Nonaka M., Shin-i T., Kohara Y. and
               Satoh N.
TITLE          Expressed genes in Ciona intestinalis (2002)
JOURNAL        Unpublished (2002)
COMMENT        Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku Kyoto, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES       source
               1..541
               /organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /clone="cibd032115"
               /tissue_type="blood cells"
               /clone_lib="Nori Satoh unpublished cDNA library, blood
               cells"

ORIGIN
Alignment Scores:
Pred. No.:      290      Length:      541
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             5        Gaps:        0

US-10-726-967A-57 (1-7) x BM037056 (1-541)

Qy      1  |||||
Db      240 GAAATTAACCTTGAACCTGAT 260

RESULT 8
LOCUS     AV982887                    567 bp      mRNA      linear      EST 15-MAR-2002
DEFINITION AV982887 Nori Satoh unpublished cDNA library, egg Ciona
ACCESSION AV982887
VERSION   AV982887.1 GI:19482908
KEYWORDS  EST.
SOURCE    Ciona intestinalis
ORGANISM  Ciona intestinalis
           Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
           Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 567)
AUTHORS   Satoh N., Satou Y., Kohara Y. and Shin-i T.
TITLE     Expressed genes in Ciona intestinalis
JOURNAL   Unpublished (2000)
COMMENT   Contact: Nori Satoh
           Department of Zoology
           Kyoto University
           Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
           Tel: 81-75-753-4081
           Fax: 81-75-705-1113
           Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES   source
           1..567
           /organism="Ciona intestinalis"
           /mol_type="mRNA"
           /db_xref="taxon:7719"
           /clone="c1eg25d24"
           /tissue_type="whole animal"
           /dev_stage="egg"
           /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
Pred. No.:      313      Length:      579
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             5        Gaps:        0

US-10-726-967A-57 (1-7) x BW212514 (1-579)

Qy      1  |||||
Db      259 GAAATTAACCTTGAACCTGAT 279

RESULT 10
LOCUS     BM192913                    612 bp      mRNA      linear      EST 05-NOV-2002
DEFINITION BM192913 Nori Satoh unpublished cDNA library, young adult Ciona
           intestinalis cDNA clone c1ad098c16 5', mRNA sequence.

```

```

/clone="c1eg25d24"
/tissue_type="whole animal"
/dev_stage="egg"
/clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
Pred. No.:      306      Length:      567
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             2        Gaps:        0

US-10-726-967A-57 (1-7) x AV982887 (1-567)

Qy      1  |||||
Db      215 GAAATTAACCTTGAACCTGAT 235

RESULT 9
LOCUS     BM212514                    579 bp      mRNA      linear      EST 05-NOV-2002
DEFINITION BM212514 Nori Satoh unpublished cDNA library, egg Ciona
           intestinalis cDNA clone c1eg070108 5', mRNA sequence.
ACCESSION BM212514
VERSION   BM212514.1 GI:24627128
KEYWORDS  EST.
SOURCE    Ciona intestinalis
ORGANISM  Ciona intestinalis
           Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
           Phlebobranchia; Clonidae; Ciona.
REFERENCE 1 (bases 1 to 579)
AUTHORS   Satou Y., Shin-i T., Kohara Y. and Satoh N.
TITLE     Expressed genes in Ciona intestinalis (2002c)
JOURNAL   Unpublished (2002)
COMMENT   Contact: Nori Satoh
           Department of Zoology
           Kyoto University
           Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
           Tel: 81-75-753-4081
           Fax: 81-75-705-1113
           Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES   source
           1..579
           /organism="Ciona intestinalis"
           /mol_type="mRNA"
           /db_xref="taxon:7719"
           /clone="c1eg070108"
           /tissue_type="whole animal"
           /dev_stage="egg"
           /clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
Pred. No.:      313      Length:      579
Score:          35.00    Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels:      0
DB:             5        Gaps:        0

US-10-726-967A-57 (1-7) x BW212514 (1-579)

Qy      1  |||||
Db      259 GAAATTAACCTTGAACCTGAT 279

RESULT 10
LOCUS     BM192913                    612 bp      mRNA      linear      EST 05-NOV-2002
DEFINITION BM192913 Nori Satoh unpublished cDNA library, young adult Ciona
           intestinalis cDNA clone c1ad098c16 5', mRNA sequence.

```


ACCESSION BM192913
VERSION BM192913.1 GI:24607315
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Ciona.
REFERENCE 1 (bases 1 to 612)
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1. .612
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad098c16"
/issue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satoh unpublished cDNA library, young adult"

ORIGIN
Alignment Scores:
Pred. No.: 333 Length: 612
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-726-967a-57 (1-7) x BM192913 (1-612)

Qy 1 G|U|I|E|A|S|N|L|E|U|C|I|U|T|H|A|P 7
|||||
Db 308 GAAATTAACCTTGAAACTGAT 328
|||||

RESULT 11
BM220061 618 bp mRNA linear EST 06-NOV-2002
LOCUS BM220061
DEFINITION Intestinalis cDNA clone cieg098e04 5', mRNA sequence.
ACCESSION BM220061
VERSION BM220061.1 GI:24736441
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Ciona.
REFERENCE 1 (bases 1 to 618)
AUTHORS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1. .618
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cieg098e04"

FEATURES
source

/tissue_type="whole animal"
/dev_stage="egg"
/clone_lib="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
Pred. No.: 337 Length: 618
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-10-726-967a-57 (1-7) x BM220061 (1-618)

Qy 1 G|U|I|E|A|S|N|L|E|U|C|I|U|T|H|A|P 7
|||||
Db 44 GAAATTAACCTTGAAACTGAT 64
|||||

RESULT 12
AV959453 638 bp mRNA linear EST 14-MAR-2002
LOCUS AV959453
DEFINITION Intestinalis cDNA clone ciad13b07 5', mRNA sequence.
ACCESSION AV959453
VERSION AV959453.1 GI:19447752
KEYWORDS EST.
SOURCE Ciona intestinalis
ORGANISM Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cloniidae; Ciona.
REFERENCE 1 (bases 1 to 638)
AUTHORS Satoh,N., Satou,Y., Kohara,Y. and Shin-I,T.
TITLE Expressed genes in Ciona intestinalis
JOURNAL Unpublished (2000)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1. .638
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="ciad13b07"
/issue_type="whole animal"
/dev_stage="young adult"
/clone_lib="Nori Satoh unpublished cDNA library, young adult"

ORIGIN
Alignment Scores:
Pred. No.: 349 Length: 638
Score: 35.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-726-967a-57 (1-7) x AV959453 (1-638)

Qy 1 G|U|I|E|A|S|N|L|E|U|C|I|U|T|H|A|P 7
|||||
Db 346 GAAATTAACCTTGAAACTGAT 366
|||||

RESULT 13
BM220124 638 bp mRNA linear EST 06-NOV-2002
LOCUS BM220124
DEFINITION Intestinalis cDNA clone cieg098h02 5', mRNA sequence.

```

ACCESSION      BM220124
VERSION        BM220124.1  GI:24736562
KEYWORDS
SOURCE         Ciona intestinalis
ORGANISM       Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE      1 (bases 1 to 638)
AUTHORS        Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE          Expressed genes in Ciona intestinalis (2002c)
JOURNAL        Unpublished (2002)
COMMENT        Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
  source
    1..638
    /organism="Ciona intestinalis"
    /mol_type="mRNA"
    /db_xref="taxon:7719"
    /clone="c1eg098h02"
    /tissue_type="whole animal"
    /dev_stage="egg"
    /clone_1lb="Nori Satoh unpublished cDNA library, egg"

ORIGIN
Alignment Scores:
  Pred. No.:      349      Length:      638
  Score:          35.00     Matches:      7
  Percent Similarity: 100.00%  Conservative: 0
  Best Local Similarity: 100.00%  Mismatches: 0
  Query Match:    100.00%  Indels:      0
  DB:             5        Gaps:          0

US-10-726-967a-57 (1-7) x BM220124 (1-638)

QY      1  GUUUAAnLeuGUThrAap 7
Db      306  GAAATTAACCTTGAAACTGAT 326

RESULT 14
LOCUS      BM208590      657 bp      mRNA      linear      EST 05-NOV-2002
DEFINITION      BM208590 Nori Satoh unpublished cDNA library, egg Ciona
                intestinalis cDNA clone c1eg101h1 5', mRNA sequence.
ACCESSION      BM208590
VERSION        BM208590.1  GI:24623204
KEYWORDS
SOURCE         Ciona intestinalis
ORGANISM       Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE      1 (bases 1 to 657)
AUTHORS        Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE          Expressed genes in Ciona intestinalis (2002c)
JOURNAL        Unpublished (2002)
COMMENT        Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
  source
    1..657
    /organism="Ciona intestinalis"
    /mol_type="mRNA"
    /db_xref="taxon:7719"
    /clone="c1eg101h1"
    /tissue_type="whole animal"

```

```

ORIGIN
Alignment Scores:
  Pred. No.:      360      Length:      657
  Score:          35.00     Matches:      7
  Percent Similarity: 100.00%  Conservative: 0
  Best Local Similarity: 100.00%  Mismatches: 0
  Query Match:    100.00%  Indels:      0
  DB:             5        Gaps:          0

US-10-726-967a-57 (1-7) x BM208590 (1-657)

QY      1  GUUUAAnLeuGUThrAap 7
Db      264  GAAATTAACCTTGAAACTGAT 284

RESULT 15
LOCUS      BM239812      684 bp      mRNA      linear      EST 07-NOV-2002
DEFINITION      BM239812 Nori Satoh unpublished cDNA library, cleaving embryo Ciona
                intestinalis cDNA clone c1c1105m03 5', mRNA sequence.
ACCESSION      BM239812
VERSION        BM239812.1  GI:24762001
KEYWORDS
SOURCE         Ciona intestinalis
ORGANISM       Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
               Phlebobranchia; Clonidae; Ciona.
REFERENCE      1 (bases 1 to 684)
AUTHORS        Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE          Expressed genes in Ciona intestinalis (2002c)
JOURNAL        Unpublished (2002)
COMMENT        Contact: Nori Satoh
               Department of Zoology
               Kyoto University
               Sakyo-ku, Kyoto 606-8502, Japan
               Tel: 81-75-753-4081
               Fax: 81-75-705-1113
               Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES
  source
    1..684
    /organism="Ciona intestinalis"
    /mol_type="mRNA"
    /db_xref="taxon:7719"
    /clone="c1c1105m03"
    /tissue_type="whole body"
    /dev_stage="cleaving embryo"
    /clone_1lb="Nori Satoh unpublished cDNA library, cleaving
                embryo"

ORIGIN
Alignment Scores:
  Pred. No.:      377      Length:      684
  Score:          35.00     Matches:      7
  Percent Similarity: 100.00%  Conservative: 0
  Best Local Similarity: 100.00%  Mismatches: 0
  Query Match:    100.00%  Indels:      0
  DB:             5        Gaps:          0

US-10-726-967a-57 (1-7) x BM239812 (1-684)

QY      1  GUUUAAnLeuGUThrAap 7
Db      217  GAAATTAACCTTGAAACTGAT 237

Search completed: August 26, 2005, 02:36:56
Job time : 2148 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 26, 2005, 01:09:57 ; Search time 139 Seconds
(without alignments)
82.402 Million cell updates/sec

Title: US-10-726-967A-57
Perfect score: 35
Sequence: 1 EINFELD 7

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 segs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlh
-CGN2_1/USPTO.spool_n/US10726967/runat_25082005_170154_306/app_query.fasta.1.199
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rnt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NCRM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10726967.CCN 1.1.69 @runat_25082005_170154_306 -NCPU=6 -ICPU=3
-NO MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	35	100.0	347	3	US-08-991-789A-225
C 2	35	100.0	347	3	US-09-062-451-225
C 3	35	100.0	347	3	US-09-598-326-225
C 4	35	100.0	347	4	US-09-289-198-225
C 5	35	100.0	347	4	US-09-429-755-225
C 6	35	100.0	347	4	US-09-699-295-225
C 7	35	100.0	384	4	US-09-270-767-13910
C 8	33	94.3	601	4	US-09-949-016-77889
C 9	33	94.3	601	4	US-09-949-016-77890
C 10	33	94.3	49818	4	US-09-949-016-12458
C 11	33	94.3	49829	4	US-09-949-016-14011
C 12	33	94.3	112219	4	US-09-949-016-12453

13	33	94.3	112222	4	US-09-949-016-14324	Sequence 14324, A
14	33	94.3	113186	4	US-09-949-016-17572	Sequence 17572, A
15	32	91.4	232	4	US-09-910-009A-70	Sequence 70, App1
16	32	91.4	235	4	US-09-910-009A-212	Sequence 212, App
17	32	91.4	238	4	US-09-910-009A-180	Sequence 180, App
18	32	91.4	238	4	US-09-910-009A-183	Sequence 183, App
19	32	91.4	241	4	US-09-910-009A-19	Sequence 19, App1
20	32	91.4	256	4	US-09-910-009A-209	Sequence 209, App
21	32	91.4	257	4	US-09-910-009A-107	Sequence 107, App
22	32	91.4	262	4	US-09-910-009A-299	Sequence 299, App
23	32	91.4	275	4	US-09-910-009A-131	Sequence 131, App
24	32	91.4	287	4	US-09-910-009A-236	Sequence 236, App
25	32	91.4	290	4	US-09-910-009A-341	Sequence 341, App
26	32	91.4	290	4	US-09-910-009A-113	Sequence 362, App
27	32	91.4	293	4	US-09-910-009A-338	Sequence 338, App
28	32	91.4	293	4	US-09-910-009A-344	Sequence 344, App
29	32	91.4	296	4	US-09-910-009A-154	Sequence 154, App
30	32	91.4	298	4	US-09-910-009A-251	Sequence 251, App
31	32	91.4	304	4	US-09-910-009A-113	Sequence 113, App
32	32	91.4	313	4	US-09-910-009A-116	Sequence 116, App
33	32	91.4	315	4	US-09-910-009A-374	Sequence 374, App
34	32	91.4	322	4	US-09-910-009A-49	Sequence 49, App1
35	32	91.4	325	4	US-09-910-009A-31	Sequence 31, App1
36	32	91.4	326	4	US-09-910-009A-34	Sequence 34, App1
37	32	91.4	331	4	US-09-910-009A-37	Sequence 37, App1
38	32	91.4	336	4	US-09-910-009A-191	Sequence 191, App
39	32	91.4	337	4	US-09-910-009A-40	Sequence 40, App1
40	32	91.4	337	4	US-09-910-009A-43	Sequence 43, App1
41	32	91.4	357	4	US-09-910-009A-167	Sequence 167, App
42	32	91.4	689	5	PCT-US95-06406A-9	Sequence 9, App1
43	32	91.4	39528	4	US-09-949-016-11790	Sequence 11790, A
44	32	91.4	39529	4	US-09-949-016-15475	Sequence 15475, A
45	31	88.6	793	1	US-08-371-764-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-08-991-789A-225/c
; Sequence 225, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; Reed, Steven G.
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 292
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Seed IP Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,789A
; FILING DATE: 11-Dec-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Poter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4800
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 225:
; SEQUENCE CHARACTERISTICS:

```

;          LENGTH: 347 base pairs
;          TYPE: nucleic acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;          SEQUENCE DESCRIPTION: SEQ ID NO: 225:
US-08-991-789A-225

Alignment Scores:
Pred. No.:          3.8          Length:          347
Score:             35.00         Matches:          7
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%   Mismatches:     0
Query Match:       100.00%      Indels:         0
DB:                3           Gaps:            0

US-10-726-967A-57 (1-7) x US-08-991-789A-225 (1-347)

QY          1 G|U|U|e|a|s|n|e|u|g|U|t|H|r|A|p 7
DB          114 GAAATTAAATTTAGAAACAGAT 94

RESULT 2
US-09-062-451-225/C
; Sequence 225, Application US/09062451
; Patent No. 6344530
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 297
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/062,451
; FILING DATE: 04-APR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.419C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-062-451-225

Alignment Scores:
Pred. No.:          3.8          Length:          347
Score:             35.00         Matches:          7
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%   Mismatches:     0
Query Match:       100.00%      Indels:         0
DB:                3           Gaps:            0

US-10-726-967A-57 (1-7) x US-09-062-451-225 (1-347)
```

```

QY          1 G|U|U|e|a|s|n|e|u|g|U|t|H|r|A|p 7
DB          114 GAAATTAAATTTAGAAACAGAT 94

RESULT 3
US-09-598-326-225/C
; Sequence 225, Application US/09598326
; Patent No. 6423496
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; NUMBER OF SEQUENCES: 247
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group PLLC
; STREET: 701 Fifth Avenue, Suite 6300
; City: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/598,326
; FILING DATE: 20-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 210121.419D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-598-326-225

Alignment Scores:
Pred. No.:          3.8          Length:          347
Score:             35.00         Matches:          7
Percent Similarity: 100.00%      Conservative:    0
Best Local Similarity: 100.00%   Mismatches:     0
Query Match:       100.00%      Indels:         0
DB:                3           Gaps:            0

US-10-726-967A-57 (1-7) x US-09-598-326-225 (1-347)

QY          1 G|U|U|e|a|s|n|e|u|g|U|t|H|r|A|p 7
DB          114 GAAATTAAATTTAGAAACAGAT 94

RESULT 4
US-09-289-198-225/C
; Sequence 225, Application US/09289198
; Patent No. 6586570
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C5
```

```
; CURRENT APPLICATION NUMBER: US/09/289,198
; CURRENT FILING DATE: 1999-04-09
; EARLIER APPLICATION NUMBER: US 09/062,451
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: US 08/991,789
; EARLIER FILING DATE: 1997-12-11
; EARLIER APPLICATION NUMBER: US 08/838,762
; EARLIER FILING DATE: 1997-04-09
; EARLIER APPLICATION NUMBER: PCT/US97/00485
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: US 08/700,014
; EARLIER FILING DATE: 1996-08-20
; EARLIER APPLICATION NUMBER: US 08/585,392
; EARLIER FILING DATE: 1996-01-01
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-289-198-225
```

```
Alignment Scores:
Pred. No.: 3.8      Length: 347
Score: 35.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 4      Gaps: 0
```

US-10-726-967A-57 (1-7) x US-09-289-198-225 (1-347)

```
Qy 1 G|U|l|e|a|s|n|l|e|u|g|l|u|t|t|a|p 7
Db 114 GAAATTATTTAGAAACAGAT 94
```

```
RESULT 5
US-09-429-755-225/c
; Sequence 225, Application US/09429755A
; Patent No. 6656480
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, Devin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-429-755-225
```

```
Alignment Scores:
Pred. No.: 3.8      Length: 347
Score: 35.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 4      Gaps: 0
```

US-10-726-967A-57 (1-7) x US-09-429-755-225 (1-347)

```
Qy 1 G|U|l|e|a|s|n|l|e|u|t|t|a|p 7
Db 114 GAAATTATTTAGAAACAGAT 94
```

```
RESULT 6
US-09-699-295-225/c
; Sequence 225, Application US/09699295
; Patent No. 6828431
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Linda B.
; APPLICANT: Dillon, Devin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.419C10
; CURRENT APPLICATION NUMBER: US/09/699,295
; CURRENT FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 326
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-699-295-225
```

```
Alignment Scores:
Pred. No.: 3.8      Length: 347
Score: 35.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 4      Gaps: 0
```

US-10-726-967A-57 (1-7) x US-09-699-295-225 (1-347)

```
Qy 1 G|U|l|e|a|s|n|l|e|u|g|l|u|t|t|a|p 7
Db 114 GAAATTATTTAGAAACAGAT 94
```

```
RESULT 7
US-09-270-767-13910/c
; Sequence 13910, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13910
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13910
```

```
Alignment Scores:
Pred. No.: 4.27      Length: 384
Score: 35.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 4      Gaps: 0
```

US-10-726-967A-57 (1-7) x US-09-270-767-13910 (1-384)

```
Qy 1 G|U|l|e|a|s|n|l|e|u|g|l|u|t|t|a|p 7
Db 93 GAAATCATTTTGAAACCGAT 73
```

```
RESULT 8
US-09-949-016-77889/c
; Sequence 77889, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77889
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77890

Alignment Scores:
Pred. No.: 22.7 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-726-967A-57 (1-7) x US-09-949-016-77889 (1-601)
QY 1 GluileasnlengluThraSp 7
Db 332 GAGTTAACTTGAGACAGAT 312

RESULT 9
US-09-949-016-77890/c
; Sequence 77890, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77890
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-77890

Alignment Scores:
Pred. No.: 22.7 Length: 601
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0
```

```
US-10-726-967A-57 (1-7) x US-09-949-016-77890 (1-601)
QY 1 GluileasnlengluThraSp 7
Db 412 GAGTTAACTTGAGACAGAT 392

RESULT 10
US-09-949-016-12458
; Sequence 12458, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12458
; LENGTH: 49818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12458

Alignment Scores:
Pred. No.: 3.58e+03 Length: 49818
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-726-967A-57 (1-7) x US-09-949-016-12458 (1-49818)
QY 1 GluileasnlengluThraSp 7
Db 24311 GAGTTAACTTGAGACAGAT 24331

RESULT 11
US-09-949-016-14011
; Sequence 14011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14011
; LENGTH: 49829
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14011

Alignment Scores:
Pred. No.: 3.58e+03 Length: 49829
Score: 33.00 Matches: 6
```

Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 4 Gaps: 0

US-10-726-967A-57 (1-7) x US-09-949-016-14011 (1-49829)

QY 1 GluIIeasnlengluThraSp 7
DB 24311 GAGTAAACTTGAGACAGAT 24331

RESULT 12

US-09-949-016-12453
; Sequence 12453, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12453
; LENGTH: 112219
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112219)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12453

Alignment Scores:

Pred. No.:	9.06e+03	Length:	112219
Score:	33.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	0
Query Match:	94.29%	Indels:	0
DB:	4	Gaps:	0

US-10-726-967A-57 (1-7) x US-09-949-016-12453 (1-112219)

QY 1 GluIIeasnlengluThraSp 7
DB 45259 GAGTCAACTTGAGACAGAT 45279

RESULT 13

US-09-949-016-14324
; Sequence 14324, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14324
; LENGTH: 112222
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(112222)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14324

Alignment Scores:

Pred. No.:	9.06e+03	Length:	112222
Score:	33.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	0
Query Match:	94.29%	Indels:	0
DB:	4	Gaps:	0

US-10-726-967A-57 (1-7) x US-09-949-016-14324 (1-112222)

QY 1 GluIIeasnlengluThraSp 7
DB 45259 GAGTCAACTTGAGACAGAT 45279

RESULT 14

US-09-949-016-17572
; Sequence 17572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17572
; LENGTH: 113186
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17572

Alignment Scores:

Pred. No.:	9.15e+03	Length:	113186
Score:	33.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	0
Query Match:	94.29%	Indels:	0
DB:	4	Gaps:	0

US-10-726-967A-57 (1-7) x US-09-949-016-17572 (1-113186)

QY 1 GluIIeasnlengluThraSp 7
DB 46245 GAGTCAACTTGAGACAGAT 46265

RESULT 15

US-09-910-009A-70
; Sequence 70, Application US/09910009A
; Patent No. 672726
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.

APPLICANT: Watkins, Maren
APPLICANT: Cruz, Lourdes J.
APPLICANT: Shon, Ki-Joon
APPLICANT: Jacobsen, Richard
APPLICANT: Jones, Robert M.
APPLICANT: Cartier, G. Edward
APPLICANT: Shen, Greg S.
APPLICANT: Wagstaff, John D.
TITLE OF INVENTION: Mu-Conopeptides
FILE REFERENCE: 2314-242
CURRENT APPLICATION NUMBER: US/09/910,009A
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 60/219,619
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: US 60/245,157
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/264,319
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: US 60/277,270
PRIOR FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 520
SOFTWARE: PatentIn version 3.0
SEQ ID NO 70
LENGTH: 232
TYPE: DNA
ORGANISM: *Comus characteristicus*
US-09-910-009A-70

Alignment Scores:
Pred. No.: 13.6
Score: 32.00
Percent Similarity: 100.00%
Best Local Similarity: 85.71%
Query Match: 91.43%
DB: 4
Length: 232
Matches: 6
Conservative: 1
Mismatch: 0
Indels: 0
Gaps: 0

US-10-726-967A-57 (1-7) x US-09-910-009A-70 (1-232)

QY 1 GlnIleAsnLeuGluThrAsp 7
|||
DB 80 GAGATCACTCTGCAGACCGAC 100

Search completed: August 26, 2005, 02:39:43
Job time : 165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: August 26, 2005, 01:11:28 / Search time 464 Seconds

(without alignments)
98.710 Million cell updates/sec

Title: US-10-726-967a-57

Perfect score: 35

Sequence: 1 EINLETD 7

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 733113 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-O=/cg2_1/USFTO_spool_1/US10726967/runat_25082005_170154_322/app_query.fasta_1.199
-DB=Published.Applications.NA -QFT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEX=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonum62
-TRANS=human40.ccd -LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10726967 @CGN 1.1 480 @runat_25082005_170154_322
-NCU=6 -ICPU=3 -NO MMAR -LARGESUBSTR -NEG SCORES=0 -WAIT -DSRBLOCK=100
-LOGLOG -DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published.Applications.NA:*

1: /cg2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cg2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cg2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cg2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cg2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cg2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cg2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cg2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cg2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cg2_6/prodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cg2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cg2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cg2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cg2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	35	100.0	347	9 US-09-810-936-225	Sequence 225, App
C 2	35	100.0	347	9 US-09-429-755-225	Sequence 225, App
C 3	35	100.0	347	9 US-09-924-400-225	Sequence 225, App
C 4	35	100.0	347	15 US-10-212-679-225	Sequence 225, App
C 5	35	100.0	347	18 US-10-079-137B-225	Sequence 225, App
C 6	34	97.1	2334	13 US-10-027-633-101658	Sequence 101658,
C 7	34	97.1	2334	17 US-10-027-633-101658	Sequence 101658,
C 8	33	94.3	42	21 US-10-726-967A-82	Sequence 82, App1
C 9	33	94.3	640	13 US-10-027-633-81947	Sequence 81947, A
C 10	33	94.3	640	13 US-10-027-633-81948	Sequence 81948, A
C 11	33	94.3	640	13 US-10-027-633-292882	Sequence 292882,
C 12	33	94.3	640	13 US-10-027-633-292883	Sequence 292883,
C 13	33	94.3	640	17 US-10-027-633-81947	Sequence 81947, A
C 14	33	94.3	640	17 US-10-027-633-81948	Sequence 81948, A
C 15	33	94.3	640	17 US-10-027-633-292882	Sequence 292882,
C 16	33	94.3	640	17 US-10-027-633-292883	Sequence 292883,
C 17	33	94.3	757	13 US-10-027-633-34874	Sequence 34874, A
C 18	33	94.3	757	13 US-10-027-633-34875	Sequence 34875, A
C 19	33	94.3	757	13 US-10-027-633-34874	Sequence 34874, A
C 20	33	94.3	757	17 US-10-027-633-34875	Sequence 34875, A
C 21	33	94.3	36303	14 US-10-152-724A-24	Sequence 24, App1
C 22	33	94.3	105413	17 US-10-427-923-3	Sequence 3, App11
C 23	32	91.4	105	20 US-10-639-016B-3	Sequence 3, App11
C 24	32	91.4	232	10 US-09-910-009A-70	Sequence 70, App1
C 25	32	91.4	232	19 US-10-828-478-70	Sequence 70, App1
C 26	32	91.4	232	10 US-09-910-009A-112	Sequence 112, App
C 27	32	91.4	235	19 US-10-828-478-212	Sequence 212, App
C 28	32	91.4	238	10 US-09-910-009A-180	Sequence 180, App
C 29	32	91.4	238	10 US-09-910-009A-183	Sequence 183, App
C 30	32	91.4	238	19 US-10-828-478-180	Sequence 180, App
C 31	32	91.4	238	19 US-10-828-478-183	Sequence 183, App
C 32	32	91.4	241	10 US-09-910-009A-19	Sequence 19, App1
C 33	32	91.4	241	19 US-10-828-478-19	Sequence 19, App1
C 34	32	91.4	256	10 US-09-910-009A-609	Sequence 209, App
C 35	32	91.4	256	19 US-10-828-478-209	Sequence 209, App
C 36	32	91.4	257	10 US-09-910-009A-107	Sequence 107, App
C 37	32	91.4	257	19 US-10-828-478-107	Sequence 107, App
C 38	32	91.4	262	10 US-09-910-009A-299	Sequence 299, App
C 39	32	91.4	262	19 US-10-828-478-299	Sequence 299, App
C 40	32	91.4	275	10 US-09-910-009A-131	Sequence 131, App
C 41	32	91.4	275	19 US-10-828-478-131	Sequence 131, App
C 42	32	91.4	287	10 US-09-910-009A-236	Sequence 236, App
C 43	32	91.4	287	19 US-10-828-478-236	Sequence 236, App
C 44	32	91.4	290	10 US-09-910-009A-341	Sequence 341, App
C 45	32	91.4	290	10 US-09-910-009A-362	Sequence 362, App

ALIGNMENTS

RESULT 1
US-09-810-936-225/c
; Sequence 225, Application US/09810936
; Patent No. US20020068285A1
; GENERAL INFORMATION:
; APPLICANT: Fruidakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John W.
; APPLICANT: Misher, Linda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

```

; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C11
; CURRENT APPLICATION NUMBER: US/09/810,936
; CURRENT FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-810-936-225

```

```

Alignment Scores:
Pred. No.: 29.4      Length: 347
Score: 35.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 9               Gaps: 0

```

US-10-726-967a-57 (1-7) x US-09-810-936-225 (1-347)

QY 1 GUILLEAENLEUGLUTHRASP 7

DB 114 GAAATTAATTAGAAACAGAT 94

```

RESULT 2
US-09-429-755-225/c
; Sequence 225, Application US/09429755A
; Patent No. US20020111467A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Smith, John M.
; APPLICANT: Reed, Steven G.
; APPLICANT: Misher, Lynda
; APPLICANT: Retter, Marc W.
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C6
; CURRENT APPLICATION NUMBER: US/09/429,755A
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-429-755-225

```

```

Alignment Scores:
Pred. No.: 29.4      Length: 347
Score: 35.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 9               Gaps: 0

```

US-10-726-967a-57 (1-7) x US-09-429-755-225 (1-347)

QY 1 GUILLEAENLEUGLUTHRASP 7

DB 114 GAAATTAATTAGAAACAGAT 94

```

RESULT 3
US-09-924-400-225/c
; Sequence 225, Application US/09924400
; Patent No. US20020165371A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misher, Lynda E.

```

```

; APPLICANT: Dillon, David C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C12
; CURRENT APPLICATION NUMBER: US/09/924,400
; CURRENT FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-924-400-225

```

```

Alignment Scores:
Pred. No.: 29.4      Length: 347
Score: 35.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 9               Gaps: 0

```

US-10-726-967a-57 (1-7) x US-09-924-400-225 (1-347)

QY 1 GUILLEAENLEUGLUTHRASP 7

DB 114 GAAATTAATTAGAAACAGAT 94

```

RESULT 4
US-10-212-679-225/c
; Sequence 225, Application US/10212679
; Publication No. US20030125536A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary
; APPLICANT: Hirsch, Shannon Kathleen
; APPLICANT: Dillon, David
; APPLICANT: Foy, Teresa
; APPLICANT: Houghton, Ray
; APPLICANT: Persing, David
; APPLICANT: Kalos, Michael
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.419C14
; CURRENT APPLICATION NUMBER: US/10/212,679
; CURRENT FILING DATE: 2002-08-02
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-679-225

```

```

Alignment Scores:
Pred. No.: 29.4      Length: 347
Score: 35.00        Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%  Indels: 0
DB: 15              Gaps: 0

```

US-10-726-967a-57 (1-7) x US-10-212-679-225 (1-347)

QY 1 GUILLEAENLEUGLUTHRASP 7

DB 114 GAAATTAATTAGAAACAGAT 94

```
RESULT 5
US-10-079-137B-225/c
; Sequence 225, Application US/10079137B
; Publication No. US20040073016A1
; GENERAL INFORMATION:
; APPLICANT: Fridakis, Tony N.
; APPLICANT: Reed, Steven G.
; APPLICANT: Smith, John M.
; APPLICANT: Misner, Lynda E.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Day, Craig H.
; APPLICANT: Li, Samuel X.
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.419C13
; CURRENT APPLICATION NUMBER: US/10/079,137B
; CURRENT FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-079-137B-225

Alignment Scores:
Pred. No.: 29, 4      Length: 347
Score: 35.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 18               Gaps: 0

US-10-726-967A-57 (1-7) x US-10-079-137B-225 (1-347)

QY 1 GtUtlEaSnLeuGlutHrAp 7
Db 114 GAAATTAATTTAGAAACCGAT 94

RESULT 6
US-10-027-632-101658
; Sequence 101658, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101658
```

```
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101658

Alignment Scores:
Pred. No.: 496      Length: 2334
Score: 34.00        Matches: 6
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 85.71%     Mismatches: 0
Query Match: 97.14%      Indels: 0
DB: 13              Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-101658 (1-2334)

QY 1 GtUtlEaSnLeuGlutHrAp 7
Db 443 GAAATAATCTGGAACCGAT 463

RESULT 7
US-10-027-632-101658
; Sequence 101658, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101658
; LENGTH: 2334
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-101658

Alignment Scores:
Pred. No.: 496      Length: 2334
Score: 34.00        Matches: 6
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 85.71%     Mismatches: 0
Query Match: 97.14%      Indels: 0
DB: 17              Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-101658 (1-2334)

QY 1 GtUtlEaSnLeuGlutHrAp 7
Db 443 GAAATAATCTGGAACCGAT 463

RESULT 8
US-10-726-967A-82/c
; Sequence 82, Application US/10726967A
; Publication No. US20050074456A1
; GENERAL INFORMATION:
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: Constructs for Homogenously Processed Preparations of Beta Site
```

```

; TITLE OF INVENTION: App-Cleaving Enzyme
; FILE REFERENCE: 2004345-0021
; CURRENT APPLICATION NUMBER: US/10/726,967A
; CURRENT FILING DATE: 2003-12-02
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Oligonucleotide to introduce BlnI site into soluble human
; OTHER INFORMATION: proBACE1
US-10-726-967A-82

Alignment Scores:
Pred. No.: 7.84 Length: 42
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 21 Gaps: 0

US-10-726-967A-57 (1-7) x US-10-726-967A-82 (1-42)

QY 1 GluIIeAsnLeuGluThrAsp 7
DB 27 GAAGTGAATCTAGAGACCGAC 7

RESULT 9
US-10-027-632-81947/c
; Sequence 81947, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81947
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-81947

Alignment Scores:
Pred. No.: 193 Length: 640
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 13 Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-81947 (1-640)

QY 1 GluIIeAsnLeuGluThrAsp 7
DB 32 GAATGAACTAGAGACTGAC 12

RESULT 10
US-10-027-632-81948/c
; Sequence 81948, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81948
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-81948

Alignment Scores:
Pred. No.: 193 Length: 640
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 13 Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-81948 (1-640)

QY 1 GluIIeAsnLeuGluThrAsp 7
DB 32 GAATGAACTAGAGACTGAC 12

RESULT 11
US-10-027-632-292882/c
; Sequence 292882, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
```

```
QY 1 GluIIeAsnLeuGluThrAsp 7
```

```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292882
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-292882

Alignment Scores:
Pred. No.: 193
Score: 33.00
Percent Similarity: 100.00%
Best Local Similarity: 85.71%
Query Match: 94.29%
DB: 13
Gaps: 0

US-10-726-967a-57 (1-7) x US-10-027-632-292882 (1-640)

Qy 1 GluileasnlengluThraap 7
Db 32 GAATGAATCTAGAGACTGAC 12

RESULT 12
US-10-027-632-292883/c
; Sequence 292883, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292883
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-292883

Alignment Scores:
Pred. No.: 193
Score: 33.00
Percent Similarity: 100.00%
Best Local Similarity: 85.71%
Query Match: 94.29%
DB: 13
Gaps: 0

US-10-726-967a-57 (1-7) x US-10-027-632-292883 (1-640)

Qy 1 GluileasnlengluThraap 7
Db 32 GAATGAATCTAGAGACTGAC 12

RESULT 13
US-10-027-632-81947/c

; Sequence 81947, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81947

Alignment Scores:
Pred. No.: 193
Score: 33.00
Percent Similarity: 100.00%
Best Local Similarity: 85.71%
Query Match: 94.29%
DB: 17
Gaps: 0

US-10-726-967a-57 (1-7) x US-10-027-632-81947 (1-640)

Qy 1 GluileasnlengluThraap 7
Db 32 GAATGAATCTAGAGACTGAC 12

RESULT 14
US-10-027-632-81948/c
; Sequence 81948, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81948

```

; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-81948

Alignment Scores:
Pred. No.: 193 Length: 640
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 17 Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-81948 (1-640)

QY 1 GtUtleAenleuGluThrAsp 7
|||::|||
DB 32 GAATTGAATCTAGAGACTGAC 12

RESULT 15
US-10-027-632-292882/c
; Sequence 292882, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 292882
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-292882

Alignment Scores:
Pred. No.: 193 Length: 640
Score: 33.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 94.29% Indels: 0
DB: 17 Gaps: 0

US-10-726-967A-57 (1-7) x US-10-027-632-292882 (1-640)

QY 1 GtUtleAenleuGluThrAsp 7
|||::|||
DB 32 GAATTGAATCTAGAGACTGAC 12

Search completed: August 26, 2005, 02:47:09
Job time : 466 secs